REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*

2. REPORT DATE
4/3/2000

3. REPORT TYPE AND DATES COVERED
Fourth Annual 2/1/99 - 1/31/2000

4. TITLE AND SUBTITLE
Training Spatial Knowledge Acquisition Using Virtual Environments.

5. FUNDING NUMBERS
N00014-96-1-0379

6. AUTHOR(S)
N.I. Durlach, T.E. Wiegand, A. Brooks, S. Madden, L. Delhorne, R.L. Garnett

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Research Laboratory of Electronics
Massachusetts Institute of Technology
77 Massachusetts Avenue
Cambridge, MA 02139

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Office of Naval Research
Ballston Center Tower One
800 North Quincy Street
Arlington, VA 22217-5660

10. SPONSORING/MONITORING AGENCY REPORT NUMBER
96PR03143-00

11. SUPPLEMENTARY NOTES
The view, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy, or decision, unless so designated by other documentation.

12a. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

This report summarizes the work done by MIT in Year 4 (1 February 1999 through 31 January 2000) of the ONR Grant N00014-96-1-0937 entitled "Training Spatial Knowledge Acquisition Using Virtual Environments." It has been prepared by Nathaniel Durlach (PI), Dr. Thomas E. von Wiegand, Andrew Brooks, Sam Madden, Lorraine Delhorne and Rebecca Lee Garnett

14. SUBJECT TERMS

15. NUMBER OF PAGES

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UL

FOURTH ANNUAL REPORT FOR:

TRAINING SPATIAL KNOWLEDGE ACQUISITION USING VIRTUAL ENVIRONMENTS

(1 February 1999 TO 31 January 2000)

Nathaniel I. Durlach, Thomas E. von Wiegand, Andrew Brooks,
Sam Madden, Lorraine Delhorne and Rebecca Lee Garnett

Massachusetts Institute of Technology
Research Laboratory of Electronics
77 MASSACHUSETTS AVENUE, ROOM 36-767
CAMBRIDGE, MASSACHUSETTS 02139
617-253-2534
durlach@mit.edu

Prepared for:
Dylan Schmorrow
COGNITIVE & NEURAL SCIENCES
OFFICE OF NAVAL RESEARCH (CODE 342)
800 N. QUINCY STREET
ARLINGTON, VA 22217-5660
703-696-4502 VOICE
703-696-8343 FAX

2 Automated VE Generation System

Development and incremental refinement has continued of the Automated VE-Generation System, which allows complex VEs to be constructed with significantly reduced manual effort from commonly available two-dimensional DXF floorplan files and mechanically acquired photorealistic texture information. A second prototype of the texture acquisition apparatus was designed and constructed, having a shorter distance of travel but more precise positioning ability. This second prototype will allow higher quality textures to be obtained in areas where fine detail is important, such as those containing textual signs and pictures. Also, its more compact envelope will allow it to acquire textures from smaller spaces, such as those associated with passage doglegs or cluttered environments. This will further assist in the primary goal of the automated VE-Generation System; to reduce the labor, time, and cost required to construct realistic VEs of complex spaces.

In order to illustrate and evaluate these savings, the Automated VE-Generation System and the second scanner prototype were used to create a new model of the 7th floor of our laboratory building, similar to the one used in the Koh experiments, but in VRML format. Quantification of the relative labor expenditure is difficult,but a time saving of at least 50% was indicated. A copy of the paper describing the automated VE-Generation System that has been submitted for publication is included with this report as an appendix.

3 World in Miniature

A large portion of our effort has been focused on the creation of a World In Miniature, or WIM, as a spatial training tool. A WIM is a supernormal artifact within a VE simulation, whereby a miniaturized representation of the entire virtual world is made available to the user as an object within the world, including the location and orientation of the user's avatar (see discussion of WIMs in Stoakley, Conway, and Pausch, 1995). Users are able to hold a motion-tracked physical representation of the WIM within their hands and move it in and out of the field of view and vary its position and orientation, in order to facilitate flexible examination, much as they would examine a 3-D map in the real world if such a map were available.

The WIM follows the user during movement through the virtual environment; the user may choose to have the WIM remain in view during movement, thus providing a real-time overview of the relative position of the avatar within the world, or may choose to alternate between hiding and showing the WIM as if storing and retrieving it from a pocket. The WIM tool that has been developed is a generic immersive simulation within DIVE, a collaborative VE system, that allows any VRML or DIVE model world to be equipped with a tracked WIM of itself and undergo real-time exploration by a user. An example shot of a user within the VE looking at the WIM is shown in Figure 1.

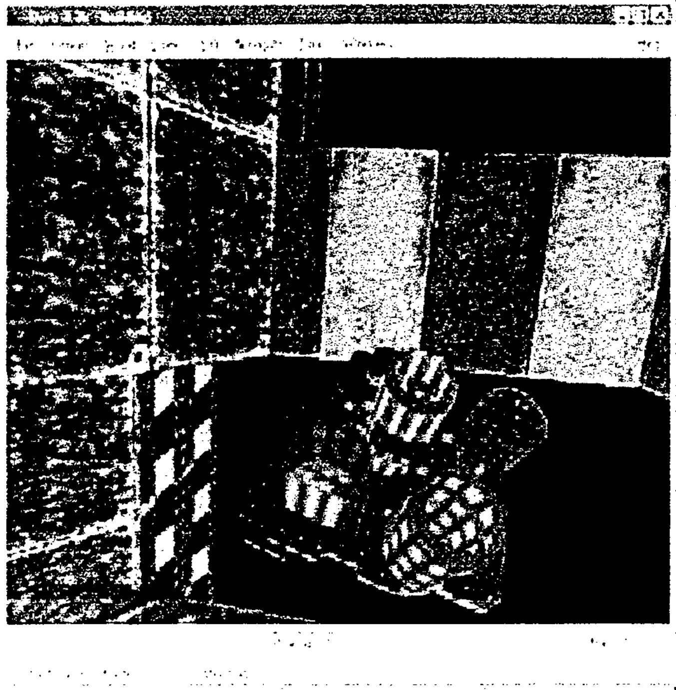

Figure 1: Example VE showing the WIM in the foreground

Figure 2: Removal of WIM detail using the saw

It is immediately apparent that a monolithic 3-D model of an indoor space would rapidly become quite useless as the level of complexity increases, due to occluding detail and the inherent difficulties in examining fine detail from an overview position. What separates our WIM from a simple virtual 3-D model of the space is that the user has the ability to disassemble it to further determine the substructure of the environment and localize the avatar. The current disassembly mode of the WIM is a transparent planar cutter or "saw" that is tracked with a hand-held physical analog just as in the case of the WIM. Wherever the saw intersects the WIM, the intersecting polygons are removed allowing the user to view the detail within. This enables the user to experience some kinesthetic feedback associated with distortion of the model, just as some kinesthetic feedback would be present when manually dissecting a real physical model. Figure 2 shows the example WIM from Figure 1 undergoing disassembly with the saw, and figure 3 shows a user immersed in the simulation, viewing the environment via a head-mounted display and operating the hand-held trackers for the WIM and the saw. Other WIM distortion modes are currently under consideration.

Figure 3: WIM simulation in sue with HMD and WIM and saw trackers

The development of the WIM will support experimentation into the usefulness of VEs in helping individuals acquire knowledge about specific spaces. Since maps are known to be an extremely useful real-world navigational aid, it seems likely that the inclusion of a three-dimensional WIM in a highly realistic VE could lead to a spatial training system that is extremely effective. The generic WIM simulator provides a platform for determining the effectiveness of such an approach to VE-assisted spatial training over a wide range of environments and a wide range of tasks to be performed in those environments.

4 References

Koh, G., von Wiegand, T.E., Garnett, R.L., Durlach, N.I., and Shinn-Cunningham, B. (1999). Use of Virtual Environments for Acquiring Configurational Knowledge about Specific Real-World Spaces: I. Preliminary Experiment. Presence 8(6).

Stoakley, R., Conway, M.J., and Pausch, R. (1995). Virtual Reality on a WIM: Interactive Worlds in Miniature. CHI'95 Mosaic of Creativity.

5 Appendices

Four documents are appended to the text of this annual report:

"Virtual Environments and the Enhancement of Spatial Behavior: A Proposed Research Agenda," by Nat Durlach, Gary Allen, Rudy Darken, Rebecca Lee Garnet, Jack Loomis, Jim Templeman, and Thomas E. von Wiegand. Under revision for *PRESENCE: Teleoperators and Virtual Environments*;

"TOADS: A Two-Dimensional Open-Ended Architectural Database System," by Samuel Madden and Thomas E. von Wiegand. Submitted to *PRESENCE: Teleoperators and Virtual Environments*;

"A Random Walk: Mastering Musketaquid," by Rebecca Lee Garnett;

"Inner Space: A Rat's Eye View," by Andrew Brooks.

Research contributing to the first two of these four papers was supported mainly by ONR grant N00014-96-1-0937 entitled "Training Spatial Knowledge Acquisition Using Virtual Environments." The third and fourth papers document unfunded participant observation studies undertaken on the personal initiative of the respective authors.

1 Executive Summary

This report summarizes the work done by MIT in Year 4 (1 February 1999 through 31 January 2000) of the ONR Grant N00014-96-1-0937 entitled "Training Spatial Knowledge Acquisition Using Virtual Environments." It has been prepared by Nathaniel Durlach (PI), Dr. Thomas E. von Wiegand, Andrew Brooks, Sam Madden, Lorraine Delhorne and Rebecca Lee Garnett

A large portion of our effort during this period has been directed towards the development of a generic simulation involving a World in Miniature, or WIM. We believe that the WIM will prove to be a valuable training aid for the acquisition of configurational knowledge of specific spaces. The generic nature of the simulation will accommodate both completely virtual environments as well as virtual copies of real-world environments, for example those created with our Automated BE Generation System, to be used in VSPAN experiments.

Additional efforts have been devoted to further improvement of the Automated VE Generation System itself, and the preparation of a paper on this system for publication. This paper is included as an appendix to this document.

A paper, "Virtual Environments and the Enhancement of Spatial Behavior: A Proposed Research Agenda" (Durlach et al., 1999) was also completed during this reporting period, marking the results of a collaborative effort with other researchers outside MIT to outline a general research and development program in the area of VE-assisted spatial training. This paper (which has been submitted for publication and is now under revision) is also included as an appendix.

In addition, as part of the work we are doing as a follow-on to our initial work on the acquisition of configurational knowledge (see Koh et al., 1999), we have initiated some basic research concerned with perceptual and cognitive factors involved in spatial updating. More specifically, this work is concerned with how errors that occur in estimating the bearing and range of targets are affected by real and imagined translations and rotations, and with the implications of these results for models of spatial perception and cognition. (Further details of this work will be presented in our next annual report.)

Finally, we include as supplementary appendices two reports of team members of personal experiences in the areas of wayfinding and acquisition of configurational knowledge that help provide insight into varying spatial problems of interest ("A Random Walk: Mastering Musketaquid," and "Inner Space: A Rat's Eye View").

Equipment acquisitions that have taken place during the year include the replacement of our aging VR4 head-mounted display (HMD) with a new Kaiser ProView 50, providing a significant boost to resolution and visual field, and the incorporation of an InterSense six-degree-of-freedom motion tracker to allow much more accurate and precise position and orientation sensing than was achievable with Polhemus equipment.

The body of this report contains more detailed descriptions of the improvements that have been made to the Automated VE-Generation System and an introduction to the WIM simulation.

ATTACHMENT NUMBER 1

REPORTS AND REPORT DISTRIBUTION

REPORT TYPES

(a) Performance (Technical) Report(s) (Include letter report(s)) Frequency: Annual

(b) Final Technical Report, issued at completion of Grant.
NOTE: Technical Reports must have a SF-298 accompanying them.

(c) Final Financial Status Report (SF 269)

(d) Final Patent Report (DD 882)

REPORTS DISTRIBUTION

ADDRESSEES	REPORT TYPES	NUMBER OF COPIES
Office of Naval Research Program Officer Terry Allard ONR 342CS Ballston Centre Tower One 800 North Quincy Street Arlington, VA 22217-5660	(a) & (b) w/(SF-298's)	3
Administrative Grants Officer OFFICE OF NAVAL RESEARCH REGIONAL OFFICE BOSTON 495 SUMMER STREET ROOM 103 BOSTON, MA 02210-2109	(c), (d) & SF-298's only for (a) & (b)	1
Director, Naval Research Laboratory Attn: Code 5227 4555 Overlook Drive Washington, DC 20375-5326	(a) & (b) w/(SF-298's)	1
Defense Technical Information Center 8725 John J. Kingman Road STE 0944 Ft. Belvoir, VA 22060-6218	(a) & (b) w/(SF-298's)	2
Office of Naval Research Attn: ONR 00CC1 Ballston Centre Tower One 800 North Quincy Street Arlington, VA 22217-5660	(d)	1

If the Program Officer directs, the Grantee shall make additional distribution of technical reports in accordance with a supplemental distribution list provided by the Program Officer. The supplemental distribution list shall not exceed 250 addresses.

TOADS: A Two-Dimensional Open-Ended Architectural Database System

Samuel Madden
(University of California, Berkeley)

&

Thomas E. von Wiegand
(Massachusetts Institute of Technology)

address correspondence to:

Dr. T. E. von Wiegand
MIT/Research Laboratory of Electronics
Fifty Vassar Street 36-755
Cambridge MA 02139

Tew@mit.edu

TOADS: A Two-Dimensional Open-Ended Architectural Database System

Abstract

The TOADS system is presented as an innovative tool for building interior-space virtual environments in two-dimensions. Existing virtual environment design tools typically operate in three-dimensions, which makes it difficult to manipulate objects on the inherently two-dimensional computer screen. TOADS allows nearly the same functionality as those three-dimensional systems in an easy-to-use two-dimensional environment. Users edit and enhance DXF floorplans with height and texture information. The software includes an inference engine which automatically identifies doors in the floorplan and generates openable polygons in the final environment. It also includes a sophisticated mechanism for embedding complex textures, such as transparent windows, at arbitrary heights in wall polygons. The entire interface is integrated with software that drives a custom texture-acquisition device. This device consists of a rack-mounted camera which captures narrow bands of textures and tiles them together to form long, continuous swaths of texture. This paper summarizes these tools and their function, and presents examples of environments which were generated with them.

1. Introduction

The Two dimensional Open-ended Architectural Database System (TOADS) is a general purpose software tool for cataloging and organizing data in three-dimensional environments. Its main goal is to manage the wealth of data which is available in architectural environments; this includes, but certainly is not limited to, floorplans, wall, ceiling, and floor textures, photographs, historical documents, locations of movable objects such as furniture, and lighting. With TOADS, users can manage these data in a coherent, easy-to-use environment, and then select relevant subsets of the data and export them to less manageable contexts, such as three-dimensional virtual environments (VRML, Inventor, etc.) or HTML-style documents containing text and pictures connected by hypertext links.

The software tools which have been created to date consist of a prototype suite which is designed to demonstrate the feasibility and usefulness of such a project. In its current incarnation, the TOADS interface consists of a 2D editor for DXF-format blueprint files, combined with several support programs that facilitate texture-capture and application. In addition to textures, these blueprints can be augmented with further information, which, in the current TOADS release, consists primarily of ceiling heights and lighting. These models can then be exported to VRML files and explored using a standard web-browser. Additional features, such as support for other input and output file formats and textual information linked to the blueprint, are not included in the initial implementation, but the software is designed with these features in mind.

1.1 State of The Art

Virtual reality has become one of the most media-hyped terms in computer science today. A rash of computer generated movies, beginning with "Toy Story" and including the more recent "Antz" and "A Bug's Life" have popularized the notion that computers are capable of generating immersive virtual universes. This is furthered by the huge number of first-person, three-dimensional video-games, among them "Doom", "Descent", and "Unreal". These applications of 3D graphics technology have made most of society aware of the possibilities that virtual reality holds. Trends in academia have drawn from popular ideas of science fiction and begun to envision an electronic world in which vast numbers of Internet users co-exist in a three-dimensional virtual community. To this end, recent literature has focused on managing complexity, bandwidth, and integration issues involved in building and rendering virtual worlds.

Researchers have taken two major approaches to this problem: the first has to do with keeping track of which objects in a world are visible in an efficient manner and spending compute time on rendering visible or probably visible objects . The second has to do with limiting the size of models by linking them together via hyperlink-style visual portals in the environments; smaller models mean faster performance, though issues of discontinuity arise when users move across the boundaries of models .

SGI, in association with a number of commercial and academic institutions, recently developed a formal specification for 3D environments called Virtual Reality Modeling Language (VRML) . Typically implemented on top of OpenGL, VRML is an OpenInventor-like language designed for making 3D environments accessible to users of desktop machines via the Internet. It is important because it provides an accepted, platform-independent, high-performance rendering engine which is extremely easy to use. As with HTML, VRML allows users with little programming experience to create dynamic, high-quality worlds which can be explored in real time. Cosmo Software, a spinoff of SGI, currently distributes CosmoPlayer, a VRML viewer for Netscape Navigator which allows Internet users to download and interact with virtual reality scenes from within their web browser.

VRML, Inventor, and other proprietary languages for describing virtual environments are not, by themselves, adequate tools for most users who wish to create virtual scenes without learning a complex programming-style language . Most end-users will not invest the time required to learn VRML, and visions of a world wide virtual environment will never be realized without user-friendly development tools. This is where TOADS fits in – as a tool to facilitate the rapid generation of high-quality, photorealistic virtual environments. Thus, TOADS, as a VRML authoring tool, is to VE design as HTML authoring tools are to web page design – it provides a non-programming based environment where most architectural VE's can be built.

1.2 More On TOADS

Previous work in the MIT RLE Sensory Communication Group suggests that manually creating photo-realistic three-dimensional environments is painstaking and tedious . This is due primarily to two concerns: first, manually describing each polygon as a sequence of numeric points that form walls, doors, windows, and ceilings is extremely slow. Once the model has been built, making changes to it is relatively difficult because of the density of the 3D data and the fundamentally 2D nature of the computer screen – even with a graphical editor, locating, selecting, and accurately moving or resizing a particular polygon can take several minutes.

The second major problem with such environments has to do with acquiring, cropping, and applying textures. Typically, textures are obtained via a digital camera, then cropped in a digital photography program to be the correct size and orientation for the polygon they're to be applied to. Then, the user is required to manually associate each polygon with one of the textures, with no way of logically applying default textures to all polygons of a certain class – e.g., making all of the doors in a building have the same appearance.

TOADS provides a solution to both of these problems. It is designed to work with existing architectural data formats (such as DXF) in which blueprints are commonly stored. This provides the user with a two-dimensional representation of the building being modeled. Furniture (and other movable objects) can be specified as two-dimensional polygonal outlines and then textured with photographs to create a realistic appearance. All objects can be readily moved and resized.

The single largest benefit of TOADS, however, lies in its ability to facilitate texture acquisition and application. In conjunction with the development of the TOADS software system, an effort is underway to build hardware devices that acquire large strips of texture by moving a camera along a section of wall (by placing the camera on either a motorized track or a semi-autonomous robot.) TOADS includes a software interface which allows it to control these devices. It also includes a texture-tiling application to precisely line up frames from the camera to create wall-sized textures without requiring any user-editing. Furthermore, because the system is intelligent about the floorplans it is working with, it can automatically detect some kinds of architectural objects (for example, doors are nearly universally represented as arcs in architectural plans) and apply intelligent texture defaults to those objects. Users can associate objects with particular categories (such as "rooms" or "windows") and apply textures to all objects in a particular grouping.

A third advantage of TOADS lies in its flexibility: it allows any geometric data with a two dimensional-representation to be edited and exported into arbitrary three-dimensional file formats. Although the system is initially directed towards manipulating blueprints, adapting it to work with geographic data (e.g. open-space or undersea environments) is as simple as writing a parser for those data. Two and three-dimensional USGS data is available for the entire world –TOADS could import and enhance such data to generate complete three-dimensional virtual environments: land could be textured, lighting effects added, and polygonal objects such as houses and trees could be drawn in.

Yet another importance of TOADS lies in its ability to handle dynamic environments. Because objects can be moved, rotated, and resized, it is easy to change the appearance of an environment without having to manually tweak polygons in three dimensions. Because TOADS saves files in a compact form, and can open and save quickly, it is possible to keep many versions of an environment around, each with slightly a different arrangement of objects, textures, and heights.

What is the importance of being able to quickly generate 3D environments? Our work is immediately heading towards experiments on the training of spatial behavior, in which human subjects are trained in virtual environments of real buildings and are then taken to the actual building and asked to perform tasks that depend on the spatial knowledge acquired by the subject in the virtual environment. Among the spaces we're considering modeling is a very cluttered warehouse; making a model of such a building by hand would be tedious and hugely time consuming, but should be considerably easier with TOADS.

On a larger scale, the system is significant in any study in which a virtual environment is used. Not only does it greatly reduce the time to build the environment, it provides a simple interface to make and catalog changes as experiments evolve.

2. Existing Tools

3D graphics tools have existed for many years, and there is such a wide variety that it is at times hard to separate one from the other. In order to properly place TOADS amongst other tools, it is important to understand how those tools work and what they do. Some mention has already been made of Inventor and VRML, but a bit more background is important to understand why some of the design decisions in TOADS were made.

2.1 DXF

DXF is the data format (front-end) that initial versions of TOADS support. DXF is AutoDesk corporation's widely used architectural and CAD drawing format. It was selected for several reasons:

- Nearly universal support: DXF support is available for many different kinds computers and platforms.

- Large existing set of data available: DXF-format blueprint files are commonly available for many different buildings. MIT provides DXF files for all of its buildings.

- Extensible: DXF provides a comment mechanism that can be used to embed TOADS specific data while still allowing other DXF parsers to view the files.

- Easy to parse: DXF is a text-based format that is very easy to parse using a simple top-down, finite-state based parser.

DXF is a two and three dimensional drawing format. Objects are separated into named layers, and object-primitives can be clustered into named blocks. The principal object primitives are lines, circles, arcs, polygons (two and three dimensional), and text. It provides a variety of drawing options which are designed with CAD and architectural needs in mind.

2.2 VRML

VRML is currently the VE file-format of choice for TOADS files. After a model has been created and enhanced with appropriate textures and ceiling heights, it is exported to the TOADS intermediate 3D format. This intermediate format can be converted to VRML using a supplied conversion tool, or to any other 3D file format using a user-provided tool and the TOADS intermediate file interfaces.

As mentioned previously, VRML was selected because it is a nearly universally supported standard for virtual environment creation. Increasingly, it is the format in which commercial

VE's are designed. Browsers are available for all platforms, and active research is being done to improve its performance and capabilities .

The VRML format offers a number of useful features: the syntax is easy to generate, consisting of text-based statements to define and transform objects; also, it is powerful, supporting fully textured environments (with transparency) with complex lighting, collision detection (in Version 2.0), and interactivity.

Interactivity is primarily accomplished via embedded JavaScripts (Netscape Corporations Java-like language originally developed to allow web-page authors to quickly insert scripts into their pages.) Objects can be linked to JavaScripts which are activated when the user approaches or clicks on them. Scripts can have a variety of effects, from initiating a simple animation to teleporting the user to a new location in the environment.

Interactivity is important to the TOADS system to allow openable objects like doors to be supported. Since Javascripts are easily embedded into VRML, TOADS can build environments with openable doors without requiring the user to compile source code.

VRML and DXF are important because they allow TOADS to operate within the domain of existing standards; users can continue to use the file formats and tools they are familiar with and still gain the benefits of an advanced, UI driven VE system like TOADS.

2.3 3D Modeling Environments

These are tools whose primary purpose is to generate three-dimensional objects or renderings of three-dimensional scenes. They often include tools to manipulate objects in three dimensions, apply lighting and textures, and generate high-quality ray-traced renderings. They differ significantly from TOADS in that, rather than concealing three dimensional details from the user, they work hard to display those details in their full glory. These are the most prevalent of three-dimensional design tools; they've existed for a number of years on all desktop platforms.

2.4 Architectural Walkthrough Programs

There are several programs designed to create 3D walkthroughs for quick prototyping of architectural environments. One of most powerful of these is Virtus Walkthrough, which presents the user with a two-dimensional overhead view of the model they are creating with a side-by-side 3D view of the environment. It allows textures to be applied to walls, lights to be specified, and a variety of polygonal objects to be defined. It can also import DXF files, and can export to VRML (version 1.0). In this respect, it is quite similar to TOADS and serves as a minimum for what TOADS should do. Despite the initial similarities, however, it is lacking in a number of areas:

- No facilities for texture management and acquisition. Walkthrough assumes that textures exist a priori and has no support for cropping and tiling textures within the program, much less an interface to hardware specifically designed to acquire textures. It does come with a large selection of pre-defined textures; unfortunately, such files are of little use when trying to design a photo-realistic environment.

- Limited expandability. Although the initial TOADS implementation may only support DXF import and VRML export, its design is such that plugging in new front and back ends is straightforward for anyone with moderate programming experience. Walkthrough is targeted at the commercial sector, and thus provides no room for programming-literate users to expand it.

- No Intelligent data management. Walkthrough provides no facilities to automatically or manually group objects beyond the creation of simple layers. Objects can't have classes (e.g. door/wall/window) and thus can't be textured or lit by object type. Even layers don't offer default textures or heights. There is no facility whereby objects can be subdivided into rooms, and the program does no intelligent processing to eliminate user-tedium (such as automatically classifying all quarter-circles as doors and automatically generating doors which can be clicked to open or close in the 3D environment.)

- Poor Performance. Walkthrough took more than 30 minutes to import a 2000 polygon DXF floorplan on a Pentium 90. TOADS opened the same file in under a minute on a comparable-speed PowerMac 7200 (PowerPC 601 at 75 Mhz). Though the Virtus software did present a 3D view of the file immediately, it is so slow at managing a large number of polygons as to be almost unusable except on the highest performance PC's. TOADS runs well on all Macintosh computers built within the last five years. The fully-textured models it exports require a speedy machine, but most of the building and design can be done on much lower performance machines.

Walkthrough does offer a 3D view of the model as it is being built. There are no plans to implement such a feature in TOADS, primarily because VRML and Inventor browsers are available for all desktop machines and can be used to prototype a model when necessary.

2.5 The Berkeley BMG System

The Building Model Generator (BMG) is a tool to automatically convert 2D floorplans into 3D environments . It does some of what TOADS does, in that its input is a floorplan and its output a 3D environment. It includes some sophisticated two-dimensional analysis which attempts to determine which areas of a model constitute rooms and corridors, and draws some of the same inferences about what kinds of lines make up windows and doors as TOADS. BMG doesn't include a user interface which allows the same sort of detailed manipulation of 2D models, and it is lacking any of the tools TOADS includes to import, manipulate, and rapidly apply 3D textures.

The BMG system is extremely interesting in that a large amount of work has been put into generating the three-dimensional models once the two-dimensional floorplan is fixed. It includes algorithms to automatically detect rooms and corridors (which TOADS will not initially include), as well as code to eliminate inconsistencies such as overlapping polygons and non-joined corners (which TOADS will only include in a rudimentary way). If possible, incorporating some of this work into TOADS would be beneficial.

Thus, the BMG system is complimentary to TOADS – it offers advanced 3D export tools with a very primitive 2D interface, while TOADS focuses on providing a coherent 2D interface for rapid environment development and places less emphasis on 3D export.

3. Design Considerations

3.1 Two-Dimensions is Intuitive

The main interface for the TOADS system is a two-dimensional, overhead view of the floorplan of the building. This is different from the conventional view of a virtual environment – most modelers present a three dimensional view. The problem with a three-dimensional interface is that the computer screen is inherently two-dimensional; thus, most interfaces are clunky and difficult to use, requiring the user to select separate tools for panning (moving in the xy plane at the current depth), zooming (changing the depth), and rotating (changing the viewing angle). This limitation makes it hard to locate specific points and place objects accurately in three-dimensions; for a complex architectural model, simply placing all the walls would require many hours. Conversely, a two dimensional interface is extremely easy to manipulate using familiar computer metaphors – the standard mouse pointer is sufficient to locate and place any line within a flat model.

Of course, there are some situations in which a three-dimensional interface is necessary; if objects exist at many different depths or have widely varying heights, there is no reasonable two-dimensional interface for model-viewing. Fortunately, in the domain of architectural spaces, most objects lie on the plane of the floor, and most of the walls have the same height. A two dimensional interface makes it extremely simple to move and place objects and walls, and navigating the model is intuitive and familiar to users of graphical user-interfaces. By allowing users to specify heights for walls and objects, with defaults that let most of the walls have the same height, little is lost in terms of functionality or realism when working in two-dimensions with architectural models.

There are some environments where it is necessary to create objects which do not lie on the floor. If these objects don't require user interaction and aren't particularly deep, they can be accurately approximated by photorealistic textures (see section 3.2 below.) There are a few conspicuous cases where these conditions can not be met; doors need to be opened and may have small thresholds such that they aren't flush with the floor; windows need to be transparent and usually don't stretch all the way from the floor to ceiling. To solve this problem, TOADS provides an interface which allows the position of transparent and openable objects to be specified within the context of a larger wall – the term "designer textures" is used to refer to such multi-piece textures.

3.2 Photorealitic vs. Truly Three Dimensional

A significant amount of time has been spent by serious computer scientists on the problem of how to manage the immense complexity which arises in large virtual environments, particularly when the entire model is represented as polygons . If every chair, computer, bookshelf and book in an environment is drawn as an individual polygon, a single floor of a small office building would require tens of thousands of different polygonal objects. This complexity is painful in two ways: first, generating such a model takes a huge amount of time, since someone has to describe the polygons for each and every object; second, the performance of rendering engines (such as VRML) falls off linearly with the number of polygons, so more polygons translates directly into worse performance.

In TOADS, we work around this problem by replacing many polygons with a single photographic-quality texture of those polygons. For example, a bookshelf with books on it is represented by a single picture rather than separate geometric objects for each book on the shelf. In many cases, it is safe to carry this process even further, so an entire wall, with bookshelves, pictures and boxes along it is represented as a single, long swath of texture. This is a standard technique for improving performance in complex virtual environments .

Of course, there are some drawbacks to this representation. Most significantly, there is no sense of depth in textures, so as a viewer gets close to a wall, all objects on it will appear to be flat; there is no motion parallax and all shadows and reflections are static. Also, there can be no interactivity in such a world – objects are fixed in their location and cannot be moved or rotated.

TOADS attempts to rectify some of these problems by allowing users to specify independent polygonal objects in the model. Of particular interest are objects with which the user interacts: doors need to be opened and closed and are thus represented as separate polygons which can be clicked on; windows are transparent and should allow the user to see into the space which is behind them. There may also be cases where an object has too much depth or is too far from a wall to be included in a flat texture. In these cases, users can specify polygons with their own textures at arbitrary locations within a model.

3.3 Texture Acquisition

Because TOADS is designed to use photo-realistic textures, a good interface for obtaining and applying the textures is necessary. Rather than requiring the user to manually take photographs, crop them in a photo-editing program, and then associate polygons with photographs, an interface to custom-designed texture acquisition hardware is built into TOADS. The hardware currently consists of a track-mounted camera which moves along a portion of wall, capturing narrow bands of texture as it moves. TOADS tiles these bands of texture into a larger texture file; because the camera is moving at a constant speed and its distance from the wall is known, it is possible to determine exactly how wide to make each band so that a seamless texture results. The scanning process is simple: the user clicks on the polygon he wishes to acquire a texture for, moves the hardware to the location he is scanning, and clicks a button to start the scanning process. TOADS automatically tiles the texture and applies it to the section of wall being scanned. The user doesn't have to worry about cropping, tiling, or scaling the images to produce photo-realistic textures.

To further generalize the texture acquisition process, the software which obtains photographs from the environment is implemented as a separate application from the main TOADS program. This allows other texture acquisition hardware to be used without modification to TOADS by developing a small interface program and using a standardized software interface.

3.4 Texture-mapped Conduits

A solution to the problem of flat texture panels which approximate walls with some depth is to introduce a mechanism that constrains the users movement and which prevents him from approaching close enough to a complex scene to perceive that it does not actually contain depth. We refer to these texture-mapped conduits as habitrails (after a concept from animal psychology referring to the paths which many animals naturally carve through their living-spaces) to indicate the areas of the model which the user is actually allowed to explore. The TOADS system allows such habitrails to be defined in a model; they are automatically set up to be transparent walls which limit the users movement.

3.5 Data Inference

The TOADS data inference engine is responsible for making decisions about how to automatically classify objects and how to generate intermediate polygons which enhance the environment without user intervention. It is an important part of the program because it significantly decreases the time a user is required to spend on tedious, repetitive tasks. Currently, it has three major components:

- Door detection and animation: Models contain doors which have a standardized appearance (a 90° arc with an edge connecting to the center of the circle the arc lies on.) TOADS automatically types such objects as doors. Furthermore, when the 3D model is created, doors are exported as flat polygons (not arcs) which have two-states: open and closed. 3D modelers which support interactivity (e.g. VRML) can link mouse click events to transitions between the door states, creating an environment in which the user can open and close doors.

- Layer to object type mappings: An interface to automatically type objects based on their layer in the DXF file is also supported. Because DXF models often separate objects into layers based on object types, it is straightforward to map all objects within a particular layer into a comparable object type.

- Automatic generation of ceiling and floor polygons: Polygonal objects, rooms, and models need to be closed objects, with tops and bottoms. Drawing these tops and bottoms by hand is slow, especially in complex models. TOADS automatically generates tight top and bottom polygons to represent the floors and ceilings of rooms and models as well as the top and bottom polygons of closed moveable objects.

There are a number of potential extensions to this system; the most important is automatic room detection, a feature present in the BMG system which may be transferable to TOADS. Rooms provide an important way of logically dividing up models and applying defaults: just as building floors have walls with similar textures, rooms are even more likely to have walls with the same texture.

3.6 Defaults and Object Property Inheritance

Modern architecture has the fortunate property that, within a particular building, almost every wall, floor, ceiling, door, and window looks the same. There are some differences due to human occupation – pictures on walls, notes taped to bookshelves, etc., but in the absence of personal touches, most architectural interiors don't include a large variety of surfaces or textures. For this reason, being able to specify model-wide defaults – default settings for every floor or wall within a model – can greatly reduce the number of individual polygons which have to be textured. For this reason, TOADS sets up texture-defaults for many common types of objects as well as a height default for the entire model. Each object in the blueprint is given a type, such as window, door, or wall, which is used to generate its texture if a specific texture is not otherwise specified.

3.7 Intermediate 3D Files

The intermediate 3D format is generated from the DXF objects which comprise the 2D model. Each object generates a 3D description of itself which is written to this intermediate file. The intermediate file can then be converted by an external tool into a specific 3D file format, such as Inventor or VRML. In this way, users can extend the capabilities of TOADS to support new 3D file-types without having to modify the source code of the main program. TOADS includes utility routines which facilitate parsing of the intermediate file, so users who need support for a particular 3D file format can easily write their own conversion tool. The intermediate format is a concise text-based description of polygons and textures which is easy to generate and parse. A tool to generate VRML-based environments from the intermediate format is provided.

4. Implementation

The TOADS Tool Suite consists of the main TOADS program plus a number of smaller programs which facilitate texture acquisition and VRML file creation. The system was built and currently runs on MacOS computers.

The TOADS program is responsible for opening DXF files and allowing the user to interact with them by adding textures, rooms, and other objects. DXF files are each displayed in their own window (Figure 1) and are manipulated via tool-palette (Figure 2). The tool palette provides tools to create rooms, polygons, habitrails, and lights, as well as basic selection and manipulation utilities.

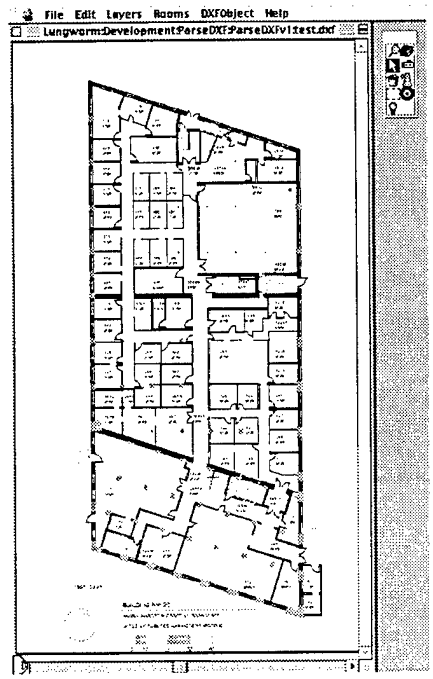

Figure 1: A screenshot of the TOADS UI with a single DXF file open.

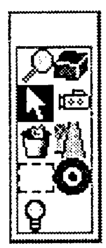

Figure 2: The Tool Palette

The three support programs are:

1) *TOADGrabber*, a tool designed to interface to the rack-mounted texture acquisition engine descibed in section 4.3. This program generates QuickTime movies which consist of the sequential frames generated as the camera moves along the wall.

2) *TOADTiler*, a tool designed to take the movies generated by the TOADGrabber program and tile them into texture images which can be directly applied to walls within the model.

3) *TOADStoVRML*, which accepts TOADS intermediate 3D files and generates VRML environments, including openable doors and transparent windows.

4.1 TOADGrabber

The TOADGrabber program serves as an interface between any Apple QuickTime compatible capture device and the TOADS environment. It captures and crops sequences of frames which are to be tiled together. It is separated from the other tools in order to isolate the system's dependence on MacOS specific technology into a single, easy replaceable package in the event the system is moved to a different platform.

4.2 TOADTiler

The TOADTiler program takes sequences of frames from TOADGrabber and tiles them into properly aligned frames. Movies from the tiler are opened and the orientation of the frames (portrait or landscape) is set up. An initial view, such as the one shown in Figure 3 is given.

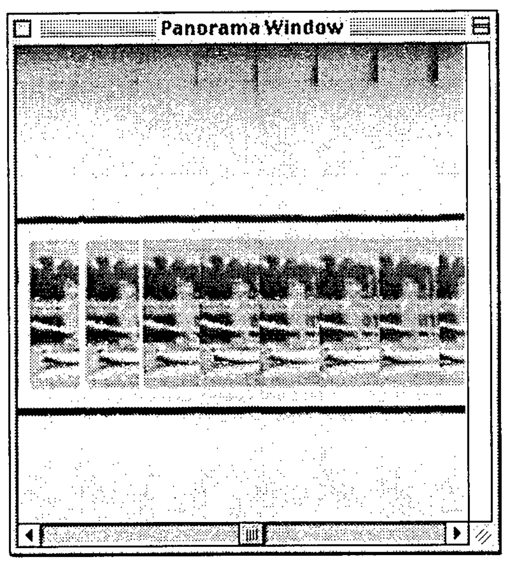

Figure 3: An improperly adjusted Panorama window.

Once the frame orientation is set-up properly, it is necessary to establish the number of pixels in each frame which equal a single movement of the camera. The "Compute Frame Size…" option from the Panorama menu provides the most intuitive way to do this; by specifying the field of view of the camera, the distance of the camera from the wall, and the number of inches in each step of the camera, TOADTiler can automatically calibrate the frame size to the proper width. Figure 4 shows this dialog.

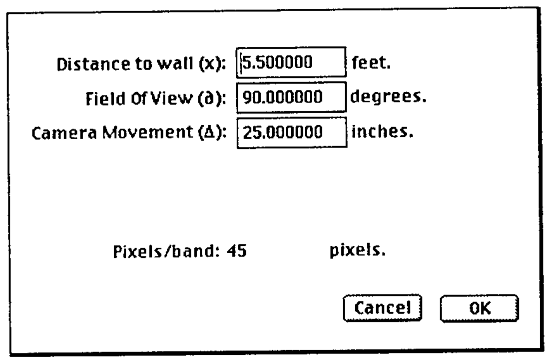

Figure 4: The Compute Frame Size Dialog.

Figure 5 shows the physical significance of each of the parameters.

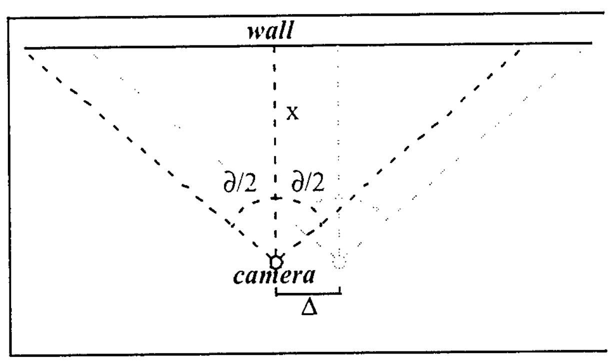

Figure 5: The x, ∂, and Δ parameters and their physical significance.

The number of pixels from each frame is computed by the program. The central region of each frame is used in order to achieve a flattened "head on" telecentric view. For the scans used in our program only the central 3 pixels are used.

Once the frame size has been properly set, images like the one in Figure 6 can be generated. Because the images are captured telecentrically, the flattened images can be applied to the walls of the model without concern that a conflicting "implied viewpoint" will confuse the viewer..

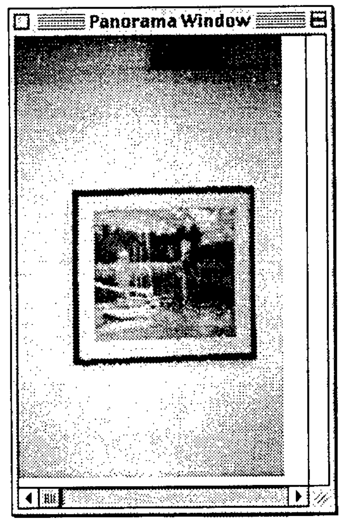

Figure 6: A properly tiled panorama.

4.2.1 Image Linearization

When scanning the ribbons of texture, we wish to capture floor to ceiling imagery. However, the required resolution at eye level (center of the vertical frame of pixels) is greater than what is required at the floor and ceiling levels. In order to achieve the desired higher pixel density in the center of the image, we utilize a wide-angle lens on the CCD camera. This lens allows the scanning rack to be placed closer to the surface being scanned while still allowing the floor to ceiling coverage. Because of the desirable compression effect at the edges ("fish eye effect") the number of pixels on the CCD which represent the central third of the image is greater than the number of pixels which represent each of the outer thirds. This effect can be seen in the calibration image reproduced in Figure 7 in which the scale markings are evenly spaced in the real world, but appear compressed toward the edges in the output from the CCD. Thus, fewer of the limited number of CCD pixels are allocated to the floor and ceiling levels of the captured image.

Calculating the inverse transform requires measuring the amount of compression. This was done by capturing an image of a vertical pole with markers every twenty inches along its length. By measuring the number of pixels between each band, it was possible to determine functional form of the nonlinearity. Figure 7 shows the captured calibration image.

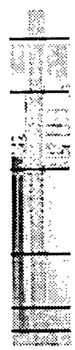

Figure 7: The calibration image. Dark lines indicate the placement of the regularly spaced markers.

By finding the offset in pixels for each marker from the center marker, it was possible to determine that the compression is linear (See Figure 8 below) with a slope of about 56 pixels per 20" (the inter-marker distance). Using this data, it is straightforward to generate the uncompressed image from the compressed pixels.

Figure 9 shows a door before and after decompression; notice that the compressed door's handle appears very far down the image, but that the decompressed door's handle in the correct position.

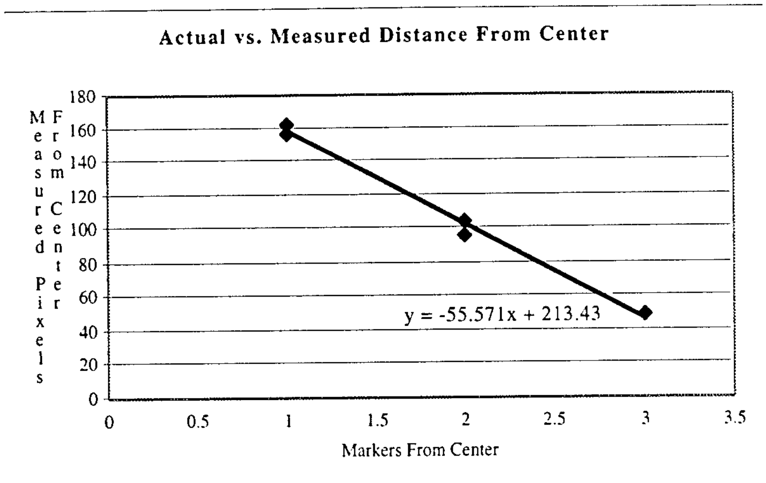

Figure 8: Graph showing the linearity of compression in the central vertical band of a captured image.

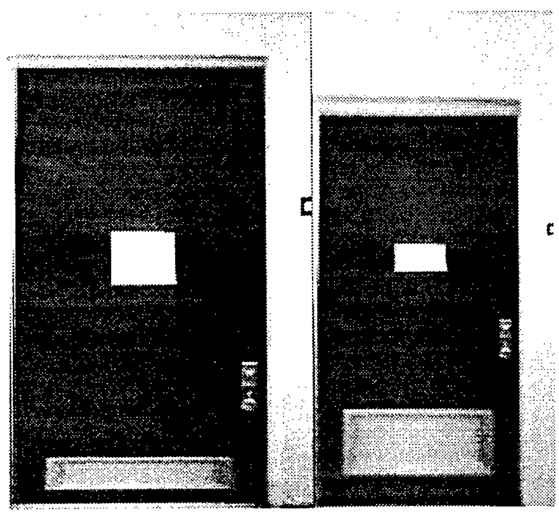

Figure 9: A door, raw input and after correction. The door handle and grating appear in the correct proportions after the corrective processing.

4.2.2 Noise Correction

TOADS employs a simple noise correction algorithm designed to eliminate bright spots in dark areas. This approach is based on the assumption that captured frames in movies have some overlap (that the frame area used for each tile in the image isn't the entire frame) and that the exact pixel offset between frames is known. Thus, corresponding areas in adjacent frames can be compared and the darkest pixels from each frame can be selected. This option is enabled via the "Noise Correction" item in the Panorama menu; the current frame is redrawn as soon as it is toggled on (or off). Figure 10 shows an image before and after noise correction; notice the white specks are no longer present (decompression has not been performed on these images).

Figure 10: An image before and after noise correction. Notice that the white dots across the top are gone, and that some irregularities in the door frame have been smoothed out.

4.3 Texture Rack

The Texture Rack is the hardware responsible for acquiring textures. It consists of a CCD camera mounted on a motorized track. The track is controlled by a stepper-motor, which allows very high-accuracy control of the position of the camera along the linear track. Each angular step of the motor moves the camera a precise distance of 0.0104 inch.

The camera provides a 320x240 pixel image. The selected lens provides a field of view of 110° enabling the camera to see the ceiling and floor in a 10-foot high space when it is just 28" from the wall. The desired compression of the image at the edges was achieved through the selection of a wide angle lens with a reasonable amount of inherent fish-eye distortion as shown in figure 11.

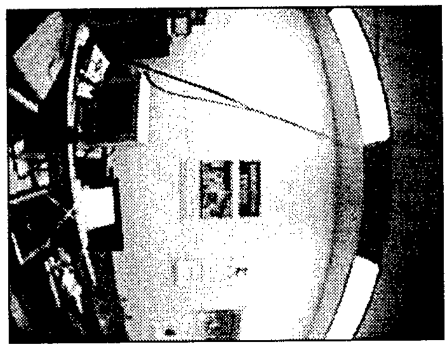

Figure 11: A sample image from the rack camera – notice the fisheye effect.

By using the stepper motor to move the camera in fine, highly controlled gradations, evenly spaced bands of pixels from the center of each position can be captured. These bands are decompressed via a simple linear transform (see section 4.2.1 above). These transformed bands can then be tiled together to form an undistorted swath of ceiling-to-floor texture having the appropriate resolution and telecentric viewpoint. (See Figure 6 for an example of a properly tiled texture.)

There are a number of issues involved in the design, implementation, and calibration of such a capture system.

4.3.1 Rack Design

Figure 12 shows the basic design of the texture rack. The belt-driven track and sled assembly was purchased as a single unit. This included the stepper motor, but none of the associated hardware or software to control it. The camera, lasers, edge sensors, controller, and battery pack were added later.

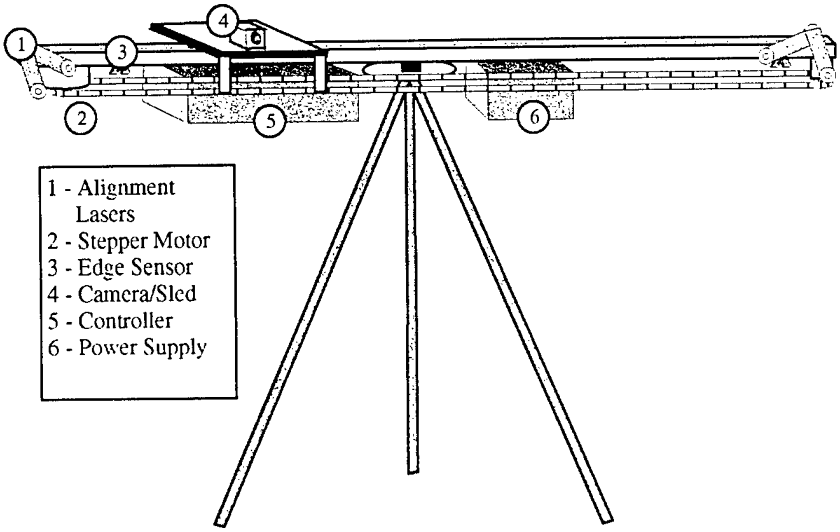

Figure 12: The Texture Rack

The lasers are used to align the rack parallel and at a fixed distance from the wall. The two outermost lasers project a beam straight ahead, and the angled lasers cross. When the beam from angled laser on the left strikes the same point as the beam from the straight-ahead laser on the right, the right edge of the rack is exactly 28" from the wall. This was done by placing the rack at this distance, moving the lasers until they were aligned, and then fixing them in place. Basic geometry guarantees that this is the only position where the lasers will align. When the angled laser on the right and the straight laser on the left are also aligned in this manner, the rack must be parallel to the wall, because both end points are equidistant.

The edge sensors determine when the camera sled has reached the left or right edge of the rack. When the rack is over them, an optical switch is closed; this switch can be read by the controller.

The battery pack uses standard Alkaline battery cells to provide a 1.5V and 12V power supply. The stepper motor is driven at 1.5V; the controller and lasers require 12V.

The controller is the heart of the texture rack. It is driven by a BASIC Stamp II (Parallax Inc.), which is a BASIC-programmable microprocessor with 16 I/O lines and a serial port. The serial port is used for communication to and from the host computer (running TOADGrabber). Using a simple set of codes, TOADGrabber instructs the texture rack to move left or right some number of steps, and can also activate the lasers and read the position of the sled. A bipolar stepper motor interface is directly implemented using MOSFET transistors connected to the I/O lines. Using a stepper motor allows the controlling program to intrinsically know the sled position (assuming the motor does not "skip" steps for some reason). Left and right limit sensors are also included so that the absolute position of the sled can be determined at startup. In addition, the positioning lasers are switched on and off under control of the BASIC Stamp (the lasers are left on except when the motor is moving).

The camera and serial port are connected to the computer via a connector on the controller box. The video output of the camera is passed directly into a video capture card on the controlling computer (the BasicStamp does not see or manipulate the video signal.)

The prototype rack is 43.5" and 4177 stepper-motor steps long. The camera is located 57" off the ground. It takes 28 seconds to move the camera from one end to the other, and 68 seconds to capture a full set of frames via TOADGrabber with 10 steps between each frame, using a PowerBook G3 with a iRez CapSure video capture card.

Figure 13 shows a photograph of the rack, with insets detailing the laser and camera assembly.

Figure 13: Photographs of the texture rack. The left inset shows the laser assembly, the right the camera and sled.

It is important to note that the texture-acquisition rack is not central to the TOADS project. Any other video input device, including a handheld camera, could be used in its place, given appropriate interface programs (like the TOADGrabber) which can process images from them. The scanning rack is an example of such a device, and was built because it produces more consistently aligned images than a handheld camera, and facilitates telecentric scanning.

4.4 3D File Export

At any point during model design, floorplans can be exported to their 3D representation. The actual export process is straightforward – each line or polygon edge is represented by a 3D polygon with the user-specified height and texture. As previously mentioned, this representation is an intermediate format specific to TOADS – it is concise and very simple to write and parse. The TOADStoVRML tool converts this intermediate format file into the corresponding VRML representation, complete with embedded JavaScripts for the interactive elements of the model.

4.5 Limitations and Future Enhancements

In the process of designing TOADS, a number of fundamental limitations and useful long-term enhancements have come to light; a few are presented here.

4.5.1 3D Export Limitations

The 3D export engine in the current TOADS systems is very primitive. It attempts to do no cleanup of the model during the export and extrusion process – this might include combining adjacent, co-linear, wall segments into a single segment or removing redundant or invisible edges from polygons. The models it generates are also completely flat – that is, all objects are grouped under one top-level VRML (or Inventor) node. If the 3D files were instead hierarchical – separated into logical groups of nearby objects – performance could be improved because most rendering engines can determine which logical groups are visible and should be drawn at any one time. This grouping would most likely be done on a room-by-room basis. Rooms can be specified in the current system, but they are not used during the export process – beginning to use them is an important future enhancement.

4.5.2 Texture Acquisition Hardware

The current system is predicated on the idea that an easy-to-use, high speed computer controlled texture acquisition device is available. The texture-rack prototype deployed for this project may not be the ideal acquisition device for several reasons. First, its relatively short length means that it doesn't capture an area much larger than a handheld camera, although it does allow more precise control over the width of the area, allows for telecentric scanning, and eliminates the need for tedious photo-retouching and manual alignment of images. Second, because it must be placed a fixed distance from the area it is acquiring, it sometimes requires shuffling of furniture and other objects to get close enough to walls to acquire them properly. The current version of the rack doesn't include options for scanning floors, ceilings, or tops of objects. Such enhancements could be carried out using multiple cameras or mirrors, with substantially the same hardware.

The initial plans for the TOADS system called for developing a semi-autonomous robot which would work in much the same way as the texture rack – by navigating parallel to walls, it could acquire and tile long bands of texture. This plan was temporarily abandoned due to time and resource constraints in developing our room scanning robot, but should be practical given continued development effort. Such a system could acquire longer bands of texture and function in smaller spaces; and ideally would also be able to make some inferences from the floorplan to locate features and automatically acquire texture for them.

Researchers around the country are working on a variety of robot-based systems which do image and texture acquisition. For example, another group at MIT, has developed a large, manned system that acquires outdoor textures from a number of static pictures of the environment. (De Couto, 1998). This system includes hardware to capture images of exterior faces of buildings. Such imagery could be fed into an enhanced version of TOADS.

4.5.3 Database Enhancements

The original vision of the TOADS system incorporated a number of different types of visual and textual information linked into floorplans; not only would environment-designers be able to feed in information about texture and ceiling heights, but they could provide infrastructure information, historical text and pictures of sites, sound and video clips, and links to HTML documents of relevance. Users would have the option of exporting end-products other than 3D environments; for example, a floorplan might export to an HTML document containing an imagemap of the floorplan where each room could be clicked on to show photographs and historical documents relating to that location. Or, additional textual information could be linked into the environment so that important pictures and text appeared in the VE when users entered a room or reached particular points-of-interest. Researchers at Columbia are currently developing augmented reality systems to allow data such as pictures, sounds, and video-clips to be overlayed on top of real-world views of scenes (Feiner, 1997). This software is based on associating those data points with a blueprint and then bringing up the appropriate data when the user is located or looking at a particular point . TOADS could be adapted to allow users to place text, video, and pictures and then export files compatible with the Columbia system.

5. Results

The goal of the TOADS project was to demonstrate the feasibility of building two-dimensional rapid-construction systems for three-dimensional virtual environments. The prototype system as it currently exists meets that goal well – several different environments from our lab at MIT have successfully been built in very short periods of time.

The first environment created was a model of the lobby of the 7^{th} floor of building 36 at MIT. This model was constructed in December 1998, before the texture acquisition hardware was complete, so textures were captured via a handheld digital camera. It took about 3 hours to capture and align the textures, and another half-hour to apply and properly set up the ceiling heights and lighting in the TOADS system. A sample screenshot is shown in Figure 14. In this image, the drinking fountain is a flat texture – notice that from a distance it is not possible to see the lack of depth. The floor and ceiling are handled through a tiled texture.

Figure 14: Screen shot of the lobby of the 7^{th} floor of building 36.

The short time to actually build the model indicates the usefulness of the TOADS tool. A past thesis (Koh 97) project in the Sensory Communication Group at MIT required one individual about two months to build a model of about half of the 7^{th} floor (approximately 5 times the number of textures as the lobby model.) This included time to acquire the textures and build the models. Based on the results from the lobby model, the same environment could be built in less than a week – the TOADS system provides an order-of-magnitude improvement in model-building time.

5.1 7th Floor Model

With the addition of real texture-acquisition hardware, it became possible to build a more complete model. Once again, the 7^{th} floor of building 36 was used, but this time the model includes the main common space of the lab – roughly four times the number of textures as the lobby model. Figure 15 shows TOADS model, with the included areas highlighted.

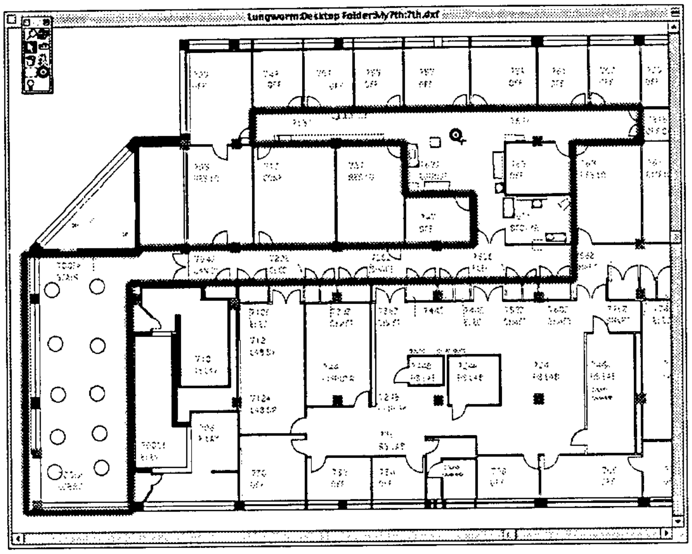

Figure 15: The Building 36 7th floor model, with fully textured areas indicated in black.

In building this model, nearly all features of the TOADS system were exercised. Large, flat texture areas were acquired using the texture rack, and smaller textures via a digital camera. Much of the furniture exists as separate polygonal models, so can be readily moved and rotated. The area outside of the lobby is set to belong to a single large room which sets up default wall, floor, and ceiling textures different from the default textures for the lobby. The windows on the lobby are transparent textures. Many of the doors and wall paintings were built using designer textures, which allowed them to be set at the proper height amongst plain-wall textures. Figures 16-18 show several different screen shots from the model.

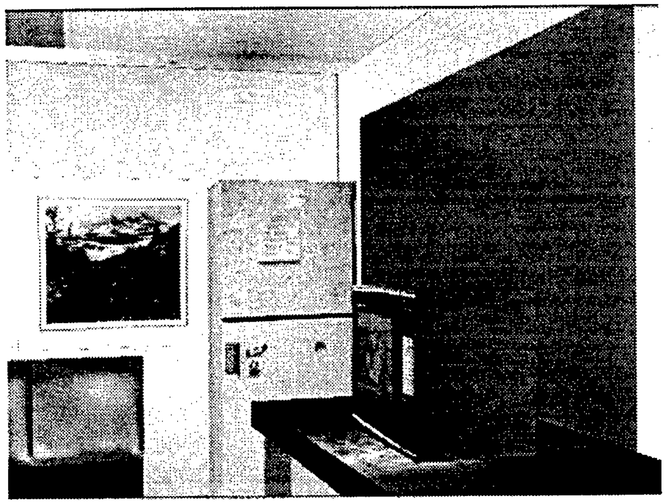

Figure 16: The fridge and microwave.

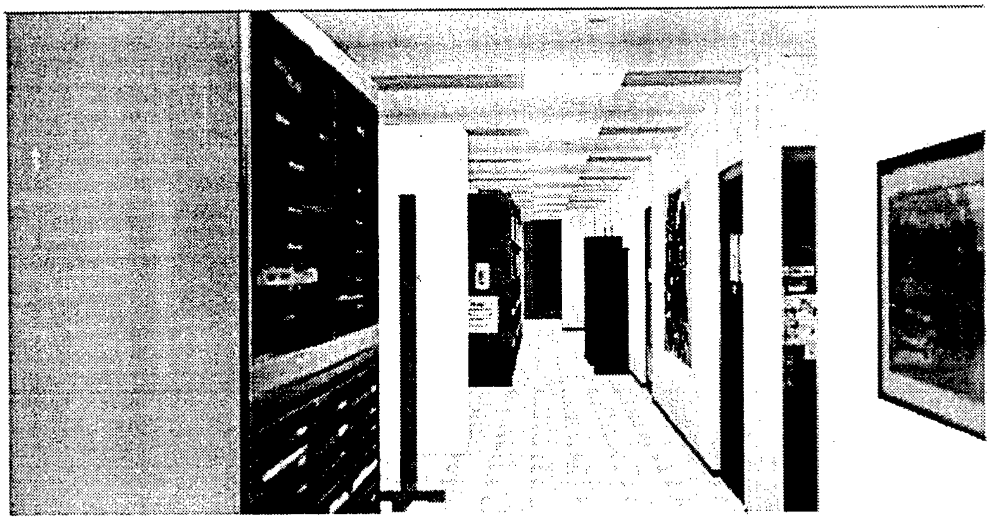

Figure 17: View past mailboxes and offices.

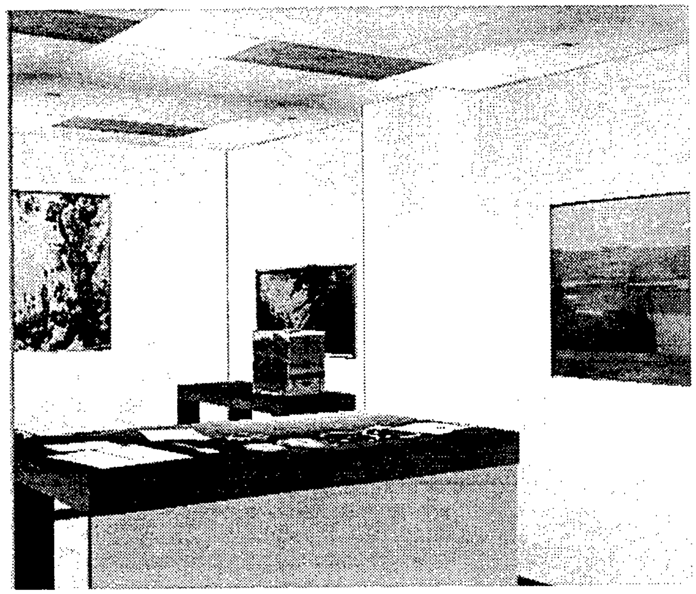

Figure 18: Secretary's desk and printer.

The bulk of the model was built in a weekend and required about 10 hours to build, including time to acquire textures. This is an extremely significant improvement over traditional architectural model design times. Furthermore, it is very easy to modify this model – if a piece of furniture needs to be moved or added, only a few seconds of editing are required.

The textured area here is comparable to the textured area in the previously mentioned (Koh 97) project, which also modeled the 7th floor of Building 36. Although building two copies of the same environment was something of a waste of resources, doing so allowed direct comparison of construction time with and without TOADS. The TOADS model has somewhat more textures and took much less time to build (ten hours versus two months. Direct comparisons of model quality are hard however, because the earlier model ran under Inventor on an SGI Onyx while the TOADS model has been run primarily on a Windows NT machine with VRML.

6. Summary

The TOADS tool suite provides a powerful software tool for designing two-dimensional interior-space virtual environments. Users can import and edit DXF models, capture and design textures, and generate VRML environments. The system in its current form is a functional software tool that has been used to generate realistic virtual environments in a fraction of the time required using previous state-of-the art tools.

Some important overall principles that guided the implementation of TOADS were presented: first, designing virtual environments in two dimensions is more intuitive and easier for users than the clumsy three-dimensional design tools normally used. Second, texture-acquisition, editing, management, and application is by far the most difficult part of virtual environment design – TOADS includes a powerful set of tools which simplify this process greatly. Finally, providing intelligent defaults for textures and ceiling heights and making inferences about doors and windows reduces the overhead and data-management burden for users.

The current TOADS system is run and built on Macintosh computers. It consists of about 20,000 lines of C++ code, much of which was designed with some consideration for portability, so that TOADS can reside on other platforms as well.

The Sensory Communication Group at MIT plans to use the TOADS system over the next few months as an integral part of its VE research, both because it greatly simplifies the process of designing and creating environments and because it provides a powerful way of organizing and maintaining the location of textures, furniture, and lighting within virtual environments.

References

Airey, J., Rohlf, J., & Brooks, F.P. (1990). Towards image realism with interactive update rates in complex virtual building environments. *ACM SIGGRAPH: Special Issue on the 1990 Symposium on Interactive 3D Graphics*, 41-50.

Autodesk, Inc. (1998). *DXF Release 14 Reference.*

Carey, Rick. (1997). *The Annotated VRML 2.0 Reference Manual.* A-W Developers Press.

Chiang, C. & Huang, A. (1997). PanoVR SD – A Software Development Kit for Integrating Photo Realistic Panoramic Images and 3-D Graphical Objects into Virtual Words, *VRST '97: ACM Symposium on Virtual Reality Software and Technology*, 147-154.

Conway, M., & Pausch, R. (1994). ALICE: A Rapid Prototyping System for Building Virtual Environments. *CHI 94: Conference Companion*, 294-295.

Coorg, S., & Teller, S. (1997). Real-Time Occlusion Culling for Models with Large Occluders. *ACM Symposium on Interactive 3D Graphics, 1997.*

Coorg, S., & Teller, S. (1996). Temporally Coherent Conservative Visibility. *ACM Symposium on Computational Geometry, 1996.*

De Couto, Douglas. (1998). *Instrumentation for Rapidly Acquiring Pose Imagery.* M.Eng Thesis, MIT.

Elvins, T. T., Nadeau, D.R.., Kirsh D., & Schul, R. (1998). Worldlets - 3D Thumbnails for 3D browsing. *CHI 98: the ACM Conference on Human Factors in Computing.*

Feiner, S., MacIntyre B., & Höllerer, T. (1997). A Touring Machine: Prototyping 3D Mobile Augmented Reality Systems for Exploring the Urban Environment. *Proceedings of the International Symposium on Wearable Computing*, 74-81.

Ginis, R. and Nadeau, D.R. (1995). Creating VRML Extensions to Support Vector Field Visualization, *VRML 95, the first annual conference on the Virtual Reality Modeling Language.*

Koh, Glenn. (1997). *Training Spatial Knowledge Acquisition Using Virtual Environments.* M.Eng Thesis, MIT.

Koh, G., von Wiegand, T.E., Garnett, R.L., Durlach, N.I., and Shinn-Cunningham, B. (1999). Use of Virtual Environments for Acquiring Configurational Knowledge about Specific Real-World Spaces: I. Preliminary Experiment. *Presence*, in press.

Lewis, Rick. (1996). *Generating Three-Dimensional Building Models from Two-Dimensional Architectural Plans.* Master's Thesis, University of California, Berkeley.

Madden, Samuel. (1999). *TOADS: A Two-Dimensional Open-ended Architectural Database System.* Master's Thesis, MIT.

Schmalstieg, D., & Schaufler, G. (1999). Sewing Worlds Together With Seams: A Mechanism to Construct Complex Virtual Environments. *PRESENCE, August 1999.*

Sharma, Rajesh. (1998). Open Inventor: A Toolkit for 3D Graphics. *The Motif Developer: July, 1998.*

von Wiegand, T. E. (1999). In B. Passero (Ed.), *Training Spatial Knowledge Acquisition using Virtual Environments* (Research Laboratory of Electronics Progress report Number 141, pp. 361-366.) Cambridge: Massachusetts Institute of Technology.

VIRTUAL ENVIRONMENTS AND THE ENHANCEMENT OF SPATIAL BEHAVIOR: A PROPOSED RESEARCH AGENDA

Nat Durlach

Research Laboratory of Electronics

Massachusetts Institute of Technology

77 Massachusetts Avenue 36-709

Cambridge MA 02139

durlach@cbgrle.mit.edu

Gary Allen

University of South Carolina

Rudy Darken

Naval Postgraduate School

Rebecca Lee Garnett

Massachusetts Institute of Technology

Jack Loomis

University of California Santa Barbara

Jim Templeman

Naval Research Laboratory

Thomas E. von Wiegand

Massachusetts Institute of Technology

VIRTUAL ENVIRONMENTS AND THE ENHANCEMENT OF SPATIAL BEHAVIOR: A PROPOSED RESEARCH AGENDA

Nat Durlach, Gary Allen, Rudy Darken, Rebecca Lee Garnett,
Jack Loomis, Jim Templeman, and Thomas E. von Wiegand

1. INTRODUCTION AND BACKGROUND

This paper is concerned with the use of virtual environment (VE) technology to train spatial behavior in the real world.

Our use of the word "behavior" here is intended to include all the processes that underlie spatial behavior, such as spatial perception, cognition, and awareness, as well as overt spatial behavior (performance) itself. In addition, however, our use of this word is intended to stress its importance; ultimately, it is the overt behavior that counts.

Rather than focusing on past results about VE-assisted training of spatial behavior, in this Forum paper we outline a research program that is aimed at achieving future advances. By doing so, we hope to stimulate thought and discussion that will lead to the realization of a comprehensive, optimally effective research program in this area. We would be delighted if readers responded by submitting their own thoughts on this topic to the forum section of PRESENCE.

Relevant general background on VEs can be found in the following books: VEs - Durlach and Mavor (1995), Barfield and Furness (1995); VE-assisted training - Seidel and Chatelier (1997); and spatial behavior - Golledge (1999). Additional important and recent material on spatial behavior can be found in Allen (1999a), Waller (1999), and Hunt and Waller (1999).

1.1 Types of Research involving Spatial Behavior and Virtual Environments (VEs)

Four kinds of research are being conducted that involve spatial behavior and VEs:

(1) VEs are being used as a research tool to help advance fundamental understanding of spatial behavior;

(2) VEs are being used to help assess spatial abilities and skills;

(3) Because VE users often find VEs confusing and difficult to navigate (people often get lost in VEs), efforts are being directed towards the development and evaluation of methods for improving spatial behavior in VEs;

(4) Research is being conducted on the use of VEs to improve spatial behavior in the real world.

Although an increasing number of articles are being published in this general area, only a few fall within Category 4.

Within Category 4, VEs can be exploited in a number of ways:

(4A) VE technology can be used as a training aid to improve spatial behavior in the real world;

(4B) VE technology can be used as a performance aid in the real world (applications of augmented reality in industrial design and medical surgery constitute important illustrations of this use);

(4C) VEs can be used as laboratories to explore and evaluate current and envisioned performance aids of essentially any type (including aids for people with sensory impairments).

In the remainder of this note, attention is focused primarily on the use of VE technology to train spatial behavior in the real world (Category 4A). Material concerned with VEs and spatial behavior can be found in Barfield and Kim (1991); Brooks (1992); Henry (1992); Darken and Sibert (1993); Bailey (1994)*; Bailey and Witmer (1994)*; May, Peruch, and Savoyant (1995); Peruch, Vercher, and Gauthier (1995); Satalich (1995); Stoakley, Conway, and Pausch (1995); Witmer, Bailey, and Knerr (1995)*; Arthur, Hancock, and Telke (1996); Aginsky, Harris, Rensink, and Beusmans (1996); Darken and Sibert (1996a, b); Ruddle, Randall, Payne, and Jones (1996); Tlauka and Wilson (1996); Witmer, Bailey, Knerr, and Parsons (1996)*;

Allen, McDonald, and Singer (1997); Bliss, Tidwell, and Guest (1997)*; Neale (1997); Ruddle, Payne, and Jones (1997); Tate, Sibert, and King (1997)*; Wilson, Foreman, Gillett, and Stanton (1997); Wilson, Foreman, and Tlauka (1997)*; Chance, Gaunet, Beall, and Loomis (1998); Colle and Reid (1998); Darken and Banker (1998)*; Goerger (1998)*; Goerger, Darken, Boyd, Gagnon, Liles, Sullivan, and Lawson (1998)*; Peterson, Wells, Furness, and Hunt (1998); Ruddle, Payne, and Jones (1998); Rossano and Moak (1998); Waller, Hunt, and Knapp (1998)*; Witmer and Kline (1998); Koh, Wiegand, Garnett, Durlach, and Shinn-Cunningham (1999)*; and Ruddle, Payne, and Jones (1999).

Many of these articles are reviewed, at least briefly, in the paper by Koh et al (1999). Of these papers, the ones that are focused on the use of VE technology for training spatial behavior in the real world are marked with an asterisk (both above and in the list of references at the end of this paper).

Some of the many general factors underlying the idea that VE technology is likely to be useful for training spatial behavior in the real world are summarized briefly in the following paragraphs.

(1) Training in the real space may be inappropriate or impossible because of inaccessibility, cost, excessive danger, security requirements, etc.

(2) A single VE installation can be used for training individuals in a wide variety of tasks in a wide variety of spaces because it is software reconfigurable.

(3) When a VE system is used for training, it is possible to automatically and reliably (a) provide immediate feedback to the trainee and (b) record the trainee's actions for later analysis.

(4) VE training systems can be used to assess basic spatial abilities and skills and to upgrade these abilities and skills, as well as to train individuals to perform specific tasks in specific spaces.

(5) VE training systems can be molded adaptively in real time to optimize the training of specific individuals at specific stages of learning under the guidance of specific instructors.

(6) VE training systems can be used for training team behavior not only by providing common, shared environments in which real team members can interact, but also by providing virtual team members with whom an individual trainee can interact.

(7) VE training systems can be used to provide unreal situations especially designed to enhance training effectiveness. For example, viewpoints, geometries, and time scales can be altered, "x-ray vision" and other "magical" navigation aids can be employed, and degrees of uncertainty can be represented. To what extent elements of this type, which are demonstrated to improve spatial behavior in the VE, will result in positive transfer to the real world, is, of course, an important topic for research. There is no reason to believe, however, that realistic VEs are the best VEs for training.

Although many of the above-mentioned virtues of VEs apply to training in general, some are specific to the training of spatial behavior. That VEs should be particularly suitable for training spatial behavior is not surprising; many of the defining characteristics of VEs (i.e., immersiveness, "3Dness") relate to spatial matters. It should also be noted that, although largely irrelevant to the central interest of this paper, one of the main strategies being used to help solve the information-overload problem is to represent the information spatially, and to exploit VE interfaces to help the user perceive, understand, and manipulate this information spatially.

1.2 Issues, Problems, and Variables Relevant to Research in Area 3A

Even within the restricted context of area 4A (i.e., ignoring areas 1, 2, 3, 4B and 4C listed above), there is a wide range of issues, problems, and variables that need to be considered. A number of the more important ones are discussed briefly in the following paragraphs.

1.2.1 Wide Range of Applications

Potentially, one can use VE technology to train specific spatial behaviors in specific spaces, to train specific spatial skills that are applicable to a wide variety of spaces, or to train improved spatial performance of a wide variety of types in a wide variety of spaces. To date, most work in the VE community has been focused on specific behaviors in specific spaces; little attention has been given to the use of VEs for training spatial behavior in general, i.e., improving a subject's general spatial abilities and skills.

Within the domain of specific applications, there is a wide variety of types of real-world spaces and types of real-world tasks for which VE-assisted training could be useful. Just to mention a few, consider a special operations team performing a rescue mission; a pilot of a remotely controlled underwater vehicle inspecting a mine field; members of a space team performing a microgravity docking mission; or a blind individual navigating a new work environment.

Even if one focused exclusively on the training of special operations teams, the range of spaces, tasks, and constraints that required consideration would be enormous. Relevant spaces would include various types of natural terrain and urban structures; and relevant methods of moving through these spaces would include walking, running, climbing, crawling, and swimming, as well as the driving of land vehicles, water vehicles, and air vehicles. There would be a need for the trainee to acquire landmark knowledge, route knowledge, and survey (configurational) knowledge, and a need to consider, in addition to many types of egocentric estimates of various spatial properties of the environment, (e.g., range and bearing to various targets in the environment), a wide variety of exocentric estimates (relating, for example, to such questions as "Is there a clear line of sight between location A and location B?", "What is the best path for person A to follow to reach position B?", or "Is person A looking at person B?"). Furthermore, the kinds of responses that would need to be considered in conducting research in this application area would include not only the slow, reflective, cognitively mediated types of responses currently the focus of interest in most VE spatial-behavior studies, but also responses that are highly reflexive and time constrained (e.g., orienting responses). In addition, it would clearly be useful to carefully monitor and assess how the trainees's attention was distributed during performance of the task. Finally, one would have to consider the various navigation-performance aids that would be available to the trainee in the real world, as well as constraints on the cost, robustness, size, and weight of any VE training systems that were to actually be fielded for the training of special operations teams.

1.2.2 Large Intersubject Differences

It is well known that basic spatial abilities and skills exhibit large intersubject differences over the population (see Waller, 1999; Allen, 1999a, b; Allen, Kirasic, Dobson, Long, and Beck, 1996; Baenninger and Newcombe (1989, 1995, 1996); Lohman, 1988; Just and Carpenter, 1985; Lorenz and Neisser, 1986; Mumaw and Pellegrino, 1984; and Loomis, Klatzky, Golledge, Cincinelli, Pellegrino, and Fry, 1993). This fact emphasizes the importance of being able to adequately assess these abilities and skills before training is initiated. Such assessment might be incorporated into a selection procedure that increased the homogeneity of those to be trained or be used to help adapt the training to the individual. In any case, such assessment could play an important role in the research on VE-assisted training of spatial behavior by reducing unexplained variation in subject performance. Currently, the variance associated with individual differences constitutes a major obstacle to research in this area. In fact, only half jokingly, it has been suggested that if one wants to be successful in this field, one has to structure one's work such that a finding of no significant difference between the performance of two systems constitutes a positive outcome. Stated differently, attempting to show that the performance of one system is significantly better than that of another is hopeless. The only thing one can do is choose two systems of radically different cost and then show that the less expensive system is more cost-effective because there is no significant difference in the performance of the two systems.

It should also be noted that large intersubject differences exist not only in basic spatial abilities and skills, but also in other factors that can influence the success of VE training for real-world spatial behavior, such as previous experience and facility in working with VE systems and in transferring what is learned in VE systems to the real world. These features, as well as others, are considered in Waller's recent work on individual differences (Waller, 1999).

1.2.3 What Should be Trained?

The question of what should be trained (i.e., what knowledge needs to be acquired by the trainee, what skills the trainee should learn, and what abilities the trainee should strengthen) is very complex because of the many types of knowledge, skills, and abilities involved and because of the lack of adequate classification schemes or taxonomies available even when the type is specified (see, for example, Allen, 1999a, b; Allen et al., 1996; Waller, 1999).

It should also be noted that there is still considerable uncertainty about what skills and abilities *can* be trained. Many of the individual differences observed on tests of spatial skills and abilities are often regarded as biologically determined and innate, or at least as relatively unmodifiable, particularly when correlated with the subject's sex (Plomin and DeFries, 1998; Imperato-McGinley, Pichardo, Gautier, Voyer, and Bryden, 1991; Kimura, 1996; Silverman and Eals, 1992; Bever, 1992; Maccoby and Jacklin, 1974; Linn and Petersen, 1985; Harris, 1978; Caplan, MacPherson, and Tobin, 1985; Geary, 1995; Landau and Spelke, 1985; Chipman, Brush, and Wilson, 1985; Boles, 1980; and McGee, 1979).

Recent discussions of the correlations between a person's sex and various spatial abilities can be found in Waller (1999) and Rivers and Barnett (1999). Whereas Waller's work is focused on individual differences in spatial abilities with restricted attention given to correlation with the subject's sex, the book by Rivers and Barnett is focused on correlations (or the lack of them) between a wide variety of characteristics and the subject's sex with restricted attention given to spatial abilities. In both cases, however, if and when a substantial sex difference in spatial abilities is found, consideration is given to possible explanations of this difference other than simply wired-in genetic ones. Also, as indicated above, both the studies by Waller (1999) and by Allen (1999a, b) include serious discussion of how performance in various spatial tests might be related to underlying component spatial abilities.

With only modest exceptions, little effort has been devoted to training basic spatial skills and abilities (as opposed to assessing these skills and abilities or to training individuals to perform specific spatial tasks in specific spaces). Also, there are data which indicate that at least some of them are at least partially modifiable by certain types of training (see, for example, Connor and Serbin, 1985, Lohman, 1988; Baenninger and Newcombe, 1989, 1995, 1996; Knott and Gaze, 1996). Furthermore, it is possible that some of the difficulties encountered in previous attempts to train basic spatial skills and abilities could be eliminated or reduced by the use of VE technology. Clearly, the extent to which these basic skills and abilities can be enhanced by training is an extremely important issue. In addition to being of great interest scientifically, to the extent that such training could be accomplished (and could be accomplished relatively quickly and without strong critical-period learning constraints), it would greatly increase the effectiveness of essentially all applications. Not only might such training increase the mean performance level, but it also might reduce the inter-individual variation.

Quite apart from issues related to the intrinsic modifiability of basic skills and abilities, there are numerous issues related to how these basic skills and abilities should be defined, classified, and measured, and how they are related to the specific, often more complex, spatial tasks of interests in the given application domains. In other words, it is not yet clear how best to break down complex tasks into underlying constituent component tasks (i.e., to perform the relevant task analysis) and how to organize, structure, and measure performance in the component basic tasks. Thus, for example, one can consider landmark knowledge, route knowledge, and survey knowledge: topological survey knowledge vs. metric survey knowledge (as well as metric knowledge for various types of metrics); wayfinding tasks of commuting, exploring, and questing; means of wayfinding such as piloting, repetition of locomotion pattern, path integration, or navigation by cognitive map; and relevant cognitive processes such as visualization, mental rotation, perspective taking, map making, visual-spatial attention and memory. Discussions of these issues can be found in Allen (1999a, b); Allen et al. (1996), and Waller (1999).

One issue of special importance concerns the relationship between the traditional tests of basic spatial skills and abilities (e.g., see Cronbach, 1970) and the applied tasks of primary interest to the VE training community. Although there are undoubtedly reasons that can be used to explain the relatively poor ability of these tests to predict performance in these tasks (see, for example, Waller, 1999; Richardson, Montello, and Hegarty, in press), one of the major reasons may be that whereas the tests appear to focus mainly on the spatial properties of objects, the tasks focus on the spatial properties of environments. It is not at all obvious that individuals who exhibit a high (low) level of spatial behavior with objects will exhibit a high (low) level of spatial behavior with environments (Allen, 1999a, b; Waller, 1999).

1.2.4 How Should it be Trained?

Even if we knew exactly what should be trained, many uncertainties would still exist about how best to achieve the training. VE technology is capable of providing many types of simulations (3D-realistic and otherwise); of scaling space and time; of emphasizing components of interest by means of selection, distortion, and repetition; of providing a wide range of exemplars to foster generalization; and of adapting in real time to the needs of particular individuals and particular stages of the learning process. Given such capabilities, it is obvious that VE technology can be extremely useful in training spatial behavior. What is not obvious, however, is what kinds of VE systems and associated procedures will be the most cost effective for specified training objectives. In the following paragraphs, we comment briefly on a few of the options available in the choice of VE systems and the choice of training procedures to be used with such systems.

VE systems vary along a wide variety of dimensions (e.g., see Barfield and Furness, 1995, and Durlach and Mavor, 1995), each of which is likely to have some effect on the cost-effectiveness of the system for spatial training. There is generally a trade-off between what the system can offer in terms of computational power and richness of the human-machine interface on the one hand, and the cost, portability, and wearability on the other hand. Probably the most important choices with respect to the human-machine interface for training of spatial behavior concern the visual display and the means provided to the user for moving through the virtual space to be learned. Visual displays may be head-mounted or offhead, and vary in resolution, field of view, color, contrast, ergonomics, etc. Current methods of controlling movement through the virtual space include manual gesture using a control glove or a joystick; articulation commands using a speech recognition system; finger walking, walking-in-place, or real walking combined with a motion-tracking system; and a wide variety of relatively complex and bulky locomotion interfaces (e.g., an omnidirectional treadmill). Some recent results on locomotion interfaces are available in Templeman, Denbrook and Sibert (1999) and Christensen, Hollerbach, Xu, and Meek (2000). Variations in auditory displays are relatively unimportant in that high fidelity and good spatialization in azimuth (lateralization) can be provided at very low cost, and attempts to provide good spatialization in elevation and distance through the use of special techniques directed at reproducing highly realistic binaural stimulation provide only marginal benefits. (Effective rendering of auditory distance requires either exploitation of a priori information on signal intensity and spectral shape or the use of supernormal distance coding). Haptic interfaces for manual sensing and manipulation that include display as well as control features, such as force feedback or heat or vibration stimuli, can often be omitted in VEs designed for spatial training. Exceptions can arise, of course, when the training involves carrying objects, or manually piloting some virtual vehicle through the virtual space, or using one's hands to sense or manipulate features of the environment such as landmarks, walls and doors.

An additional set of factors that are important in considering the use of VE systems for training spatial behavior in specific spaces concerns the way in which the space is represented in the simulation. Beyond issues related to the fidelity and realism of the visual images, there are fundamental issues related to the viewpoints and scales associated with the images, and the extent to which and manner in which such factors can be easily manipulated by the user. Surprisingly, most studies concerned with VE-assisted training of spatial behavior for specific spaces have confined the subject's experiences to egocentric walk-through-type simulations, top-down views, and/or the viewing of more-or-less traditional 2D maps. Although the use of traditional 2D maps is clearly appreciated, the use of virtual 3D maps (or WIMs--Worlds in Miniature--see Stoakley et al, 1995), and their integration with "VE walkthroughs," has not yet been seriously exploited. In our opinion, the underlying reason for this neglect is excessive concern in the VE community with realism. In most cases, realism is relinquished only for the purpose of reducing cost. It is not generally appreciated that realism has fundamental limitations in VE-assisted training of spatial behavior. In general, going back and forth easily and quickly between exocentric views of maps and models of spaces and egocentric views derived from being in the space (real or virtual) can serve as a powerful tool not only for remaining oriented in a complex space, but for truly "learning the space."

In addition to providing a VE system with the appropriate characteristics and features, it is of course essential that effective training procedures be employed. Although spatial behavior and training are topics that have received considerable attention for many years, current knowledge of how best to improve a subjects' spatial behavior is relatively limited. Perhaps the training associated with the sport of orienteering and with various types of military operations best represent current knowledge in this area (see, for example, Department of the Army, 1969 and 1987; Lowry and Sidney, 1989; Paul, 1991).

For our purposes, it is useful to classify types of spatial training according to (a) whether the objective is to improve basic spatial skills and abilities or to improve specific spatial behaviors in specific spaces and (b) the extent to which the training method requires understanding of spatial behavior and of how to improve it on the part of the trainer.

Most work on the training of spatial behavior using VE has been directed towards training for specific spaces; the use of VE for training basic spatial abilities, or even for measuring such abilities, has hardly been considered. (An interesting example of the use of VE for assessing abilities can be found in Rizzo, Buckwalter, Neumann, Kesselman, Thiebaux, Larson, and van Rooyen, 1998, and in Larson, Rizzo, Buckwalter, van Rooyen, Kratz, Neumann, Kesselman, Thiebaux, and van der Zaag, 1999; see also the "Walkthrough test" in Waller, 1999). Although, as indicated above, it is often assumed that such abilities are innate and not susceptible to improvement by training, we have not yet found much evidence to justify this assumption. In fact, we suspect that a well-developed training program that exploits VE technology could greatly improve at least certain components of the complex set of abilities known as "basic spatial abilities."

One aspect of performance in tasks associated with learning an environment (such as making egocentric estimates of direction and distance to a hidden landmark) that often appears quite pronounced is response bias. On the whole, however, relatively little attention has been given to improving spatial behavior by eliminating response bias. Those investigators primarily interested in spatial behavior are concerned with determining, understanding, and modeling these biases. Those interested in VE training have not yet seriously addressed the bias question (in fact, in many studies only the absolute errors are presented and the effects of response bias and response variability are never even separated out). Whereas some biases appear to be quite general, others appear to depend on the specific trainee tested and still others on the specific features of the space in which the estimates are being made (see, for example, Koh et al., 1999, and Colle and Reid, 1998). Clearly, eliminating those response biases by training that depend on specific spatial features constitutes the greatest challenge; training to eliminate this type of bias would be greatly facilitated by having available a simple and accurate model that describes the manner in which and extent to which the bias depends on the specific spatial features.

Undoubtedly, understanding spatial behavior would be useful for designing a program to improve it; however, such understanding is not a prerequisite. For example in the training of spatial behavior in specific spaces, the benefits of familiarizing oneself with the space by means of "free exploration" can occur without the trainer understanding how these benefits are generated. Similarly, in the training of all kinds of spatial behavior, as in the training of behavior in general, performance can be greatly improved by providing immediate correct-response feedback. Such feedback will almost always reduce response bias; and sometimes it can reduce response variability as well. Just as in the use of correct-response feedback in the training of artificial neural nets to perform some classification task, good performance can be achieved without the trainer's knowledge of the differential characteristics of the classes that underlie the good classification performance. In other words, the building of the knowledge in the mind of the trainee that is necessary for good spatial performance by the trainee does not necessarily require that this knowledge be explicitly available to the trainer. All the trainer need know for feedback training is how to identify correct performance and how to provide effective feedback.

We do not mean to imply here that understanding spatial behavior is irrelevant to the training of spatial behavior, only that lack of knowledge in this area does not necessarily preclude successful training. Improved understanding of spatial behavior would not only prove useful in training applications (e.g., by enabling us to make an intelligent tutoring system designed to train spatial behavior more intelligently), but it would obviously be of great value in its own right. The fact that so little is still understood about the difficulties one has in compensating for rotations as opposed to translations, particularly when such transformations are imagined (e.g., see Rieser, 1989; Presson and Montello,1994), provides merely one illustration of how limited our knowledge is in this area.

1.2.5 Modeling the Spaces

Although we believe that VE-assisted training for spatial behavior in specific spaces can be developed that will surpass the training that can be achieved using the real spaces, there is no doubt that a strong motivation for the use of VE-assisted training is the inability to access the real space. Furthermore, when the real space is not accessible, the information that is available for modeling the space may be quite fragmented. For example, in a special-operations application, the available information might consist of a satellite photograph, an old "normal" photograph, and a verbal description by someone who had seen the place many years ago. In such situations, creating the VE model may be very difficult. Also, it may be necessary in the VE model to display indications of the confidence with which various aspects of the space have been modeled.

Even when the space is accessible, if it is very complex, constructing the model may be extremely time-consuming. Substantial simplifications may be required, and it may be necessary to indicate in the VE the kinds of simplifications that have been made.

In general, these kinds of problems, together with the underlying problem of modeling the space in such a way that positive transfer occurs, constitute a serious design challenge to those interested in building VEs for spatial training. In order to meet this challenge, it will probably be necessary to develop modeling approaches that not only provide means for indicating degrees of uncertainty and/or simplification, but also exploit to the extent possible general functional and organizational principles that can help fill in (essentially automatically) relevant detail when such detail is either absent in the basic data or eliminated for the sake of simplicity.

For example, a VE-modeling system directed at training spatial behavior in human-made spaces might incorporate a pre-existing knowledge base covering relevant architectural conventions (related to layout, placement of windows, symmetry, structural supports), cultural conventions (related to placement and functionality of rooms), and infrastructure standards (related to plumbing, electricity, telephone, climate control). A system so equipped would be able to fill-in missing information to some extent, without requiring creative intervention by the human designer.

A "smart" system of this type, which begins with a base of "common sense" background information, and which accepts both quantitative spatial input and qualitative descriptive input, would also be well-suited for integration with an intelligent tutoring system (ITS) to provide various display and training augmentations. Aside from providing indications of confidence with which various aspects of the space have been modeled, it could (together with the ITS) automatically generate training interventions related to landmarks and routes based on the functionality of the represented objects and spaces. It would also be possible to incorporate some degree of interaction between the trainee and the space, such as the ability to query the knowledge base, or facilitate exercises to strengthen perception of the space (reveal alignment of walls along structural members, reveal location of communication network boxes, reveal lighting direction and cast shadows as a function of time of day, reveal structural vulnerabilities for ordnance placement, etc.).

In general, of course, the modeling system and the ITS would have to be designed and integrated in a manner that was tailored specifically to train the real-world spatial tasks of interest.

Some material relevant to the development of "smart" VE construction systems can be found in Lewis (1996); Dorsey and McMillan (1998), Madden and Wiegand (1999), and the RLE Annual Report (1998), as well as the web sites
http://www/cs/berkeley.edu:80/~rickl/BMG/,
http://www.lcs.mit.edu/research/projects/project?name=9919,
http://www.lcs.mit.edu/research/projects/project?name=9946, and
http://graphics.lcs.mit.edu/index.html

1.2.6 Comparisons Among Different Training Systems

As would be expected, a substantial amount of current research on VE-assisted training of spatial behavior, is directed towards comparative evaluation of different systems (e.g., different visual displays or different navigational interfaces) and/or different training procedures (e.g., see Waller, 1999). This kind of research is motivated in part by a desire to increase basic understanding of how various components or features of the systems and procedures influence training outcomes and in part by the practical goal of optimizing cost-effectiveness.

Unfortunately, however, the results obtained in this experimental research often have less value than one would like because they are so hard to generalize. There is very little understanding of how the results depend on factors other than those that have been deliberately varied or on the specific tasks that have been employed in conducting the tests. Essentially all results obtained in this area need to be qualified by an almost endless list of detailed experimental conditions. Relevant theory is generally not adequate to enable one to know which of these conditions are central to the results obtained and which are peripheral.

This kind of problem is, of course, not unique to the case of VE-assisted spatial training. For example as discussed by Lintern (1996), research on flight training simulators has suffered from this same problem for the last twenty years.

1.2.7 Simulator Sickness

The problem of simulator sickness is, of course, relevant to any envisioned application of VE technology (see, for example, the discussion of cybersickness in the review of human-factor issues by Stanney, Mourant, and Kennedy, 1998). It is likely to be particularly relevant to the training of spatial behavior, however, both because of the emphasis on spatial behavior and because of the emphasis on training.

With respect to the training of *spatial behavior,* the presumed benefits of enabling the subject to move through the space, of providing the subject with a wide field of view, and of accepting fast changes in imagery (associated either with rapid egocentric movement or with major transformations of viewpoint) are all likely to be curtailed to at least some degree by their causative relationship to simulator sickness.

With respect to the *training* of spatial behavior, simulator sickness may be of special importance because of the length of time required in the VE to complete a substantial training program and because of the tendency of some subjects under certain conditions to become increasingly sensitized to the stimulus conditions responsible for causing the sickness.

1.2.8 Need for Improved Communication

There are at least three major communities of people that have interests and knowledge relevant to VE-assisted training of spatial behavior in the real world.

(1) Those focused on the development and use of VE systems (computer, electrical, and mechanical engineers);

(2) Those focused on the study of spatial behavior (psychologists and geographic information system specialists);

(3) Those focused on training (psychologists and teachers).

Unfortunately, the communication among these communities is generally rather poor. The first group has relatively limited understanding of both spatial behavior and the transfer of VE experiences to real world behavior (training transfer). The second knows relatively little about VE systems or about training. And the third group, although it is beginning to become involved in the use of VE systems for training a variety of skills and behaviors, knows relatively little about spatial behavior.

The recent ONR-sponsored workshop on VE and spatial behavior at the Naval Postgraduate School in Monterey (contact R. Darken at NPS for information on this workshop) constitutes a significant exception to this isolation; this workshop has been instrumental in improving communication among at least the first two of these groups.

2 TOWARDS AN R&D PROGRAM IN AREA 4A

2.1 Preliminary Remarks

There is no doubt that for any specific space, a virtual environment can be constructed and a VE-assisted training procedure developed that will result in spatial behavior in that specific space that is superior to what it would have been if the VE-assisted training had not taken place.

This result follows not only from a variety of empirical results reported in the literature, but also from simple logic; it is difficult to imagine how it could be otherwise. Furthermore, we believe that such benefits are likely to occur even if spatial performance aids are available in the real environment, although the magnitude of such benefits may decrease as the effectiveness of these aids increase.

The basic problem to be addressed is not whether virtual environments can be useful, but how they can be made most cost-effective. Among the questions that need to be answered are: what kinds of virtual environments should be constructed?, how can they be constructed easily and rapidly?, what kinds of training experiences should the subjects have in these virtual environments?, how should the environments and training experiences differ for different kinds of subjects?, how should one assess the basic spatial abilities of different subjects and to what extent can those abilities be enhanced by special training?, etc. etc. Obviously, to some extent, the answers to such questions will depend on the specific space and the specific spatial task for which the subject is being trained. However, it is equally true that to some extent they will be space and task independent. It should also be noted that training subjects for improved spatial behavior in general and training them for enhanced spatial behavior of a specific type in a specific space are related in two ways: Improving a subject's general spatial ability should improve spatial performance of specific types in specific spaces, and, alternatively, improving spatial performance of specific types in specific spaces should enhance one's spatial behavior in general.

2.2 Desirable Features of a VE Training System for Training Spatial Behavior

In this section, we review some of the properties of a VE training system that we believe are highly desirable, without modifications or compromises imposed by cost constraints. In other words, in this section, we focus on the effectiveness component of the ultimate cost-effectiveness goal.

2.2.1 High Fidelity, Realistic, Egocentric, Spatial Exploration Mode

The system should be capable of simulating in detail and with substantial realism the experience of being in and moving through the space in question. In this mode, the viewpoints available should be the same (egocentric) viewpoints as those that would be available in the real space, and the method (or methods) of movement in the VE should simulate as closely as possible the real ones (e.g., walking through the space, climbing over obstacles, driving through the space, etc.). In addition, the VE should include simulations of the performance aids that would be available in the real situation being simulated.

2.2.2 Supernormal Features

Although it is obvious that the realistic simulation mode would be useful for training (for the same reasons that access to the real space being simulated would be useful), it is equally obvious that the system should include modes that incorporate a variety of supernormal features. One such feature would involve control of object transparency so that one could look through objects and see what was behind them. An associated feature would provide the ability to move through such objects. A further set of features concern viewpoint and scale. The system should include miniature 3D models of the space (Worlds in Miniature - WIMs) with controllable viewpoints and scales, and provisions for building and altering such models as well as viewing them from different angles and with different zoom factors. It is already known that traditional 2D maps and traditional top-down views in VE systems, both of which are extremely weak versions of 3D miniature models (WIMs), have positive transfer effects. Furthermore, once such models are introduced, it provides a base for adding further useful features, such as immediately understandable you-are-here and facing-in-this direction indicators and movement control schemes.

An additional set of supernormal features that should be available concerns the methods that are available to the user for moving through the space. Although most of the unrealistic movement methods employed in the past have been chosen for cost or convenience reasons, it is not at all clear that such methods are always inferior. Moving by "magic carpet," or changing the position of an avatar in a miniature model may, for certain purposes, be much more efficient. Even under conditions in which it is beneficial to use a movement method that involves a sense of effort or mechanical work done, it is possible that an unrealistic method would be appropriate (e.g., "finger walking" in place of real walking).

Closely related to the use of unrealistic methods of movement within and through a space are the transitions back and forth between the realistic-space mode and the WIM mode. These transitions need to be structured in a manner that enables the user to control location, viewpoint, and scale at will, but within constraints designed to minimize disorientation.

Further special features that might be incorporated include those related to the suppression of simulator sickness. Such features might be included not only in connection with transitions from realistic-space mode to WIM mode, but also in connection with other possibly disturbing changes in imagery. For example, preliminary work has indicated that simulator sickness associated with relatively large and rapid head movements in head-mounted displays having a large field of view can be reduced by blanking out the visual display during these movements (DiZio and Lackner, 1998).

In general, of course, the value of a particular supernormal feature will depend on the extent to which it leads to positive transfer.

2.2.3 Active Training Procedures

The system should be capable of providing an active training program, not only the opportunity for free exploration of the environment (or for the scientific study of untrained spatial behavior). Although the specific training program that is most appropriate will depend on the specific task for which the subject is being trained (and also, perhaps, on the abilities of the specific subject), it is obvious that the system will have to include facilities for measuring subject responses and providing feedback to the subject about these responses that will enable the subject to reduce the magnitude and/or frequency of response errors. Because much of the previous work on spatial behavior involving VEs has been directed towards characterizing the subjects' behavior rather than improving it, relatively little attention has been given to the use of such feedback. We would expect that such feedback would be very effective in reducing response bias; the extent to which it can be used to reduce response variability is less clear. To the extent that it is realizable, the system should include an intelligent tutoring system (ITS) that provides the trainee with insight into the sources of the errors that are made and with guidelines for the development of improved mental models.

2.2.4 System Flexibility

The system should also be capable of addressing a wide variety of tasks in a wide variety of VEs. Such capability would not only ensure the usefulness of the system for a wide variety of specific projects, but also for enhancing the general spatial abilities of a given subject. In addition, it would facilitate study of transfer from one VE to another, and one task to another (as well as transfer from VE to real world). The tasks should include landmark recognition tasks, route-following tasks, and a wide variety of configurational knowledge tasks, as well as a variety of response modes.

2.2.5 Efficient VE Construction

The system should be capable of providing a wide variety of VEs, associated training procedures, and experimental formats, rapidly and easily. The greater the time required and the greater the expertise required to achieve various configurations, the less useful the overall installation will be.

2.2.6 Assessment and Enhancement of Basic Spatial Skills and Abilities (BSSAs)

Finally, the system should include programs for assessing the basic spatial skills and abilities of experimental subjects. Improved knowledge of such skills and abilities would enable one to (a) improve experimental studies by reducing the unexplained and uncontrolled variance in the data and (b) improve training by adapting the training program to the needs of the individual subject. Also, to the extent possible, the system should include programs for strengthening these basic spatial skills and abilities.

2.3 Cost-Effectiveness Trade-Offs

It is relatively easy to specify desirable features in a VE system for training spatial behavior. It is considerably more difficult to determine the cost-benefit of such features.

Perhaps the cost-benefit issues that have received the greatest attention to date concern the value of developing VE training systems that provide "realistic, high-fidelity" simulations. In the area of spatial behavior, the issues of primary concern have been related to computational power, quality of the visual display, and character of the mobility interface. Auditory displays are generally inexpensive and haptic interfaces have often been regarded as irrelevant for the training of spatial behavior. With regard to computational power and visual displays, issues such as frame rate, stereoscopy, time delay, field of view, visual detail and resolution have been central. System costs related to visual imagery in a VE system vary over almost 3 orders of magnitude (with the highest cost currently being approximately $1M for a complete high-end "CAVE" system). At present, there are no data available that are adequate to estimate the increased benefit in training spatial behavior associated with increased fidelity/resolution of visual imagery and an increased field of view (associated, in turn, with increased cost of the visual-display component of the VE training system).

The situation in the area of mobility interfaces is similar in the sense that a wide variety of such interfaces are now becoming available that cover a wide range of approximations to realism and that have a wide range of associated costs (in this case, the cost range is roughly 5 orders of magnitude with the maximum again being roughly $1M). The situation is also similar in that if one confines one's attention solely to spatial behavior, the cost-benefit picture for increased realism is unclear. In the event that the training is intended to include elements related to specific body postures, positions, and actions (as would be the case in many military training situations) as well as to spatial behavior, then, of course, the cost-benefit picture for mobility interfaces is radically modified.

The overall cost-benefit picture becomes even less clear when one includes the idea that fidelity/realism is not really the goal anyway, and that optimal VE-assisted training will undoubtedly involve intentional deviations from realism to provide superior training performance. Obviously, analysis of cost-effectiveness will need to include studies of the effects of various types of supernormal features as well as the effects of various levels of realism. The situation is further complicated by the requirement for some applications that the systems have a small "footprint." And, as always, attention will need to be given not only to how changes in the training system influence performance in the VE, but how they influence subsequent performance in the real world (i.e., how they affect training transfer).

2.4 Schematic Diagram of Program Elements

The accompanying figure provides a schematic diagram of the elements involved in the creation and use of a VE-assisted training system for enhancing spatial behavior.

[FIGURE ONE ABOUT HERE]

Figure 1. Schematic diagram of program elements.

The box on the left labeled VE-Assisted General Subject Treatment is intended to include both the assessment of subjects' basic spatial skills and abilities and, to the extent possible, the strengthening of these skills and abilities through training that is not tied directly to a specific task or specific space.

The General Knowledge ellipse at the top is intended to include all general knowledge about human perception, cognition, and performance, about training, and about VE systems, that is relevant to spatial behavior and to the design of facilities and procedures to be used in the VE-Assisted Training for specific Tasks and Spaces, as well as to the design of the general subject treatment functions just mentioned. To date, almost all research concerned with the application of VE technology to the training of spatial behavior in the real world falls under this heading. Also included are studies directed towards the basic understanding of spatial behavior (in either real or virtual environments) rather than on applications to training. Analysis of spatial-behavior errors and construction of theories of spatial processing and mental model formation that predict such errors is obviously a major concern in this domain.

The Specific Info ellipse is intended to include all those elements that need to be added (i.e., that cannot be specified in advance) to a general VE training system in order to achieve a system for training the specific task/space in question. In many cases, the information that is available will be fragmentary and heterogeneous, and methods must be developed for processing this information and incorporating it in the training system in a manner that is both efficient and effective.

The three facilities hexagrams are intended to include general VE facilities needed for a training system that is to be used for a wide variety of training tasks (computational facilities, head-mounted displays, etc.); special facilities associated with specific tasks that are not likely to be included in the basic facilities (such as radiant-heat and odor interfaces for training firefighters or special hardware and software for gathering and inputting visual images of particular spaces); and facilities for assessing and improving a subject's general spatial skills and abilities (such as a VE mental-rotation test for subject assessment).

2.5 Proposed Projects

2.5.1 Comprehensive Report and Annual Workshops

The area considered in this white paper is both very complex and highly interdisciplinary. In order to integrate relevant knowledge in a fashion that is optimally useful to the pursuit of the stated goals, it is recommended that (a) a comprehensive report be prepared describing the state of the art in this area and (2) annual workshops be held to discuss crucial topics in this area. Both the preparation of the report and the workshops should involve representatives of the communities associated with (1) Virtual Environments, (2) Spatial Behavior, (3) Training, and (4) Specific Application Domains. Careful thought must be given to the prioritization of the various possible application domains.

2.5.2 General Subject Treatment

2.5.2.1 Matrix of Spatial Tasks versus Basic Spatial Skills and Abilities (BSSAs)

A theoretical effort should be made to generate a comprehensive matrix of spatial tasks versus the spatial skills and abilities required to perform these tasks effectively, with emphasis on spatial tasks that are relevant to specific application domains. The spatial-task coordinate of the matrix should include important variables in the environment (man-made urban, natural terrain, undersea, zero gravity, etc.) and in the means available for moving through the environment (walking, driving a vehicle, piloting a ship or plane, etc.), as well as in the goals of the task. Once such a matrix were available, it could then be used, together with the extensive results of past efforts to identify basic skills and abilities in the psychological literature, to determine what we want to mean by "basic spatial skills and abilities" (BSSAs).

2.5.2.2 VE-Assisted Assessment of BSSAs

Based on the results of the theoretical analysis referred to in 2.5.2.1, a study should be conducted to determine appropriate methods for evaluating an individual's BSSAs. Unlike traditional methods used to evaluate BSSAs (many of which focus on spatial properties of manipulable objects rather than of environments; and almost all of which are confined to the use of paper-and-pencil tests), these new methods should exploit evolving VE technology. Inasmuch as one of the main defining characteristics of VE is interaction in a representation of 3D space, VE is a natural tool to be used for this kind of assessment (and such use constitutes a major opportunity to demonstrate the usefulness of VE in the spatial domain).

2.5.2.3 Modeling Human Performance in Exercising BSSAs

In addition to determining these assessment methods, efforts should be made to characterize and model human performance in the exercise of the BSSAs. Such research is important not only to advance general understanding of spatial behavior, but also to support attempts to strengthen the BSSAs via training (as considered in 2.5.2.4 below). One problem of particular interest in this modeling area concerns the difficulties many individuals have imagining or adapting to the effects of rotations.

2.5.2.4 Tests of Models - Interindividual

Once significant progress has been made on projects 2.5.2.1, 2.5.2.2, and 2.5.2.3, it will be necessary to conduct an extensive series of experiments to check the accuracy of the predictions derived from the results of these projects. More specifically, it will be important to evaluate our ability to predict performance on relevant spatial tasks from measurements of BSSAs on an individual-by-individual basis.

2.5.2.5 VE-Assisted Training of BSSAs

If the results of project 2.5.2.4 are sufficiently encouraging, it will then be important to determine (a) the VE-assisted training techniques that are most efficient for strengthening the BSSAs, and (b) the extent to which the application of such optimum techniques actually enhances the BSSAs. Again, as in the *assessment* of BSSAs, it seems clear that VE has great potential for *training* BSSAs. To the extent that such training is successful, it will have great practical significance. Furthermore, whatever the outcome, the results of such studies are likely to have important implications for theories addressed to the origin and susceptibility-to-change of various BSSAs.

2.5.2.6 *Tests of Models - Intraindividual*

If project 2.5.2.5 is successful, i.e., we can improve BSSAs in individuals by training, it will then be important to conduct the longitudinal, intraindividual counterpart of the experiments conducted in project 2.5.2.4, i.e., see if performance in applied spatial tasks of interest improves in the manner predicted from the improvements in the BSSAs measured in the individual subjects.

2.5.2.7 *Development of Prototype General-Subject-Treatment System*

Finally, once the BSSAs have been identified and methods for measuring them have been developed, and it has been demonstrated on both an individual basis and an intraindividual basis that performance on applied spatial tasks can be predicted from measurements of BSSAs, and the manner in which BSSAs can be strengthened by training has been explored, it will be appropriate to develop a prototype General-Subject-Treatment System that can serve as a preprocessor to the system labeled "VE-Assisted Training for Specific Tasks in Specific Spaces" in the schematic diagram shown in Sec. 2.4. Such a system is likely to be worthwhile developing even if BSSAs are found to be essentially untrainable; the system's assessment function all by itself should prove extremely valuable.

2.5.3 Training for Specific Tasks/Spaces

Ideally, all the individuals to which this training for specific tasks/spaces will be addressed will have been previously assessed and trained in general spatial behavior as outlined in the schematic diagram shown in Sec. 2.4 and as defined and developed according to the results of the research discussed in Sec. 2.5.2 above.

2.5.3.1 Development of Realistic Simulations

Although realistic simulations may not be optimum for training purposes, attempting to create them is important for a number of reasons. First, such simulations can play an important role in the development of baseline trainers. The fact that realistic simulations may not be optimal for training does not imply that they will not be useful for training. The more realistic the simulations and the methods of interaction, the less likely it is that the training transfer will be negative. Second, under certain conditions, they can serve as a substitute for the real world in tests of training transfer for studies in which the training makes use of unrealistic simulations (either supernormal or degraded). Finally, attempting to develop realistic simulations constitutes an important and useful challenge to system designers; the degree of success in constructing realistic spatial simulations can serve as a measure of the skills and technologies available for constructing spatial simulations in general. Although the tools that need to be developed to construct realistic simulations are not identical to those needed to develop supernormal simulations, there is clearly a large overlap.

2.5.3.2 Development of Effective Training Systems

Along with strengthening our ability to create realistic simulations of specific spaces, efforts need to be directed towards the development of effective systems for helping people "learn" these spaces.

Three components or features of such systems that are likely to prove important for training spatial behavior and about which research is badly needed are Locomotion Interfaces (LIs), Worlds in Miniature (WIMs), and Intelligent Tutoring Systems (ITSs).

<u>Locomotion Interfaces (LIs)</u> are only one of many types of interfaces that need to be improved to enhance spatial training. However, they deserve special attention in the envisioned program because of their unique relationship to spatial training. Whereas improvements in other components of the human-computer interface, such as visual displays, are also needed to

enhance spatial behavior, they do not have to be driven by this need; they will be driven by needs associated with other (more commercially potent) applications.

Although it is intuitively reasonable to think that how one moves through the space to be learned can have significant impact on how well the space is learned, relevant results in this area are still limited. The large body of results available in the literature dealing with the differential effects of active vs. passive movement are generally inadequate to determine the relative merits of the different methods for moving through virtual space (hand gestures, joystick movements, speech commands, mechanical interfaces for leg movements) with respect to learning the space and training transfer. In general, work in this area is needed both to develop locomotion interfaces and to evaluate them on a comparative basis for learning to perform a variety of specific tasks in a variety of specific spaces.

Worlds in Miniature (WIMs), which are small, 3D, manipulable replicas of the VE space within the VE, provide a major opportunity to create more effective VE training systems. Like a map, they enable the user to obtain a controllable exocentric view of the space which the user is attempting to move through or to learn. However, unlike a map, they can be 3D and disassembled (to enable the user to see the spaces inside the 3D structure). Also, a replica of the user can be located in the WIM to provide a natural you-are-here-and-facing-this-way indicator. In addition, movement through the space can be achieved by moving the replica of the user in the WIM.

Based on what is already well established about the usefulness of maps, as well as on simple logic, it is obvious that WIMs have great potential not only for aiding the user to navigate about in the VE (and prevent getting lost in the VE), but also to gain configurational knowledge about the space and to contribute strongly to positive training transfer to the real world. Focal points of required research in this area relate to details of WIM construction and user manipulation of the WIM, and to when and how the user should move back and forth between exocentric and egocentric views of the space.

Intelligent Tutoring Systems (ITSs) have received substantial attention in the general VE training area; however, relatively little research has been directed towards the use of ITSs for training spatial behavior. This is due perhaps to the fact that activities related to spatial learning are so common; unlike learning how to operate a relatively novel complex system (e.g., a control panel for a new nuclear plant), almost all individuals have had substantial experience learning unfamiliar spaces. Thus, in studying peoples' spatial behavior, or even in training them to learn a new space, there is great reliance on relatively free exploration and relatively little attention paid to determining optimum types of feedback and to the development of an insightful tutoring system. Although there is no guarantee that ITSs will make a large contribution here, it seems clear that attempts should be made to develop appropriate ITSs and to evaluate the extent to which they accelerate the desired training. Even if such attempts are not sufficiently successful to culminate in a good ITS, what is learned in these attempts should prove valuable for enhancing the effectiveness of human instruction in this area and/or computer-assisted training. In general, work in this area can have significant payoff prior to the completion of a theoretical model that adequately describes the trainee's state as a function of time and that can serve as a core component of a complete ITS. To a large extent of course, our ability to develop effective ITSs or improved human instruction programs will depend on how well we can understand and model spatial behavior.

Finally, it is important to note that despite the general recognition that positive transfer to the real world is what counts if one's goal is improved spatial behavior in the real world, issues related to transfer are often ignored. In particular, there appears to be a strong tendency when considering various types of artificial training augmentations in VEs to focus almost exclusively on whether or not the augmentation is likely to improve spatial behavior in the VE; relatively little consideration is given to whether or not it will improve spatial behavior in the real world when the augmentation is not available. Clearly, serious thought needs to be given to the construction of a model that will enable investigators to improve their ability to predict which training augmentations that are realizable and useful in the VE but not realizable in the real world (e.g., as performance aids) will have positive transfer effects.

2.5.3.3 Degradation Experiments

In 2.5.3.2, we discussed some of the research thrusts that are important to pursue in the attempt to develop an effective VE system to train spatial behavior in specific tasks/spaces.

Given that such a system can be developed (a result about which we have considerable confidence), it is then important to determine the extent to which, and the manner in which, cost-saving alterations of the training system degrade training performance. In other words, one must examine cost-effectiveness trade-offs.

In general, we see two complementary ways of conducting such research. First, and most obvious, one can take the systems and procedures that have been shown to provide good training and measure how training is degraded as the system and associated procedures are modified to reduce costs. Because the range of possible application conditions is large, the goal of this effort should be to produce cost-effectiveness curves, not merely isolated points. Among the major components that need to be examined in these degradation studies are the visual display, the locomotion interface, and the overall power of the computer system.

Second, the above experimental research on degradation of the VE training system should be accompanied by experimental research on degradation of the normal real-world training configuration. For example, important insight into the effects of employing different amounts of visual resolution and field of view could be obtained by degrading normal visual perception in real-world training through the use of head-mounted blinders and diffusing screens. Similarly, it would be possible to get insight into the effects of altering locomotion interfaces by comparing training performance in the real world under different movement conditions (e.g., walking, driving a golf cart, being transported in a wheel chair). Although certain experimental results on the effects of passive vs. active movement are available in the literature, many of these results are not relevant to acquisition of spatial knowledge or to training transfer.

2.5.3.4 Development of a "Smart" VE-Assisted Training System

A smart VE-assisted training system requires not only a VE system, but also an intelligent tutoring system (ITS). Furthermore, in at least some cases (such as a special operations mission under severe time constraints), it requires a method for constructing the appropriate VE and ITS quickly and from incomplete information. This problem, i.e., the problem of constructing the VE and ITS quickly and from incomplete information in a manner that is effective for training the specific real-world task in question, is extremely challenging.

Solving this problem requires the development of a high-level (meta) system that (1) incorporates a general knowledge base about spaces, spatial behavior, and spatial training (including what is known about training transfer from synthetic worlds to real worlds); (2) has inputs for receiving whatever information is available about the specific space and task in question; and (3) has outputs for guiding the construction of the VE and ITS that are most effective for the specific training required.

Clearly, the field of VE-assisted spatial training is not sufficiently advanced to realize such a meta-system at this time. However, it is not too early to begin work on some of the components. For example, one can visualize the development of a subsystem concerned with the construction of simulated architectural or building spaces that is designed to make use of severely degraded drawings and floor plans, aerial photos, verbal descriptions, and knowledge about such factors as the function of the building, the geographical area in which it is located, the time at which it was built, and the construction engineering practices prevalent in the culture at that place and time. By appropriate integration, interpolation and extrapolation of these kinds of data, together with indications in the resulting VE showing the degree of uncertainty associated with various characteristics and elements of the building, the meta-system could construct a VE that, although not perfectly realistic or accurate, was highly useful.

Ideally, of course, in constructing the VE, the meta-system would also have to take account of how the resulting VE (together with the ITS the meta-system would construct), could best be used to achieve the specified training goal. Different goals would lead to different decisions by the meta-system on how best to integrate and exploit all the information provided by the database and the various types of inputs.

2.5.3.5 *The Training of Teams*

In a sense, team training can be regarded as a special case of training individuals, and essentially all of the above remarks apply to training teams as well as individuals. Even though a team is involved, it is the individuals who make up the team that must be trained. Also, of course, to the extent that team members can be adequately represented by virtual people (with large and appropriate behavioral repertoires) in the VE used for training the individual, only a single VE station is required.

However, for many purposes, and probably for many years, effective team training will require active participation by multiple human users at multiple VE stations or in a single grand station using a complex projection system and multiple trackers. In addition, a focus on training a *team* of individuals (the main focus of most military training programs) leads to consideration of many factors that would be ignored without this focus. In general, team training must address the skills associated with planning and coordinating the placement, movement, and actions of groups in rapidly changing spatial environments, and support the training of groups of individuals to function as an integral unit. Awareness of the spatial locations and orientations of other team members is likely to be crucial in team training. Also crucial perhaps is the ability to estimate accurately the mental models of spatial locations and orientations that exist in the brains of these other team members. These kinds of higher-order spatial behaviors, combined with the added complexity resulting from the existence possibly of a large number of team members and of rapid changes in the tactical situation and in the spatial configuration of the team, make the spatial training of teams a particularly challenging area and one in which the power and flexibility of VE-assisted training systems is likely to prove particularly useful.

3. CONCLUDING REMARKS

The material presented in Sec. 2.5 above outlines briefly a variety of projects that we believe should be considered in constructing a serious program for R&D in the area of VE-assisted training of spatial behavior.

Although some of these projects appear relatively advanced (VE-assisted training of BSSAs - 2.5.2.5; development of an effective ITS - 2.5.3.2; development of a "smart" VE-assisted training system - 2.5.3.4; and the training of teams - 2.5.3.5), we nevertheless have included them because we believe they are exceedingly important to the ultimate realization of the goal and should be pursued in parallel with the other projects.

As indicated in the Introduction, we strongly urge readers of PRESENCE interested in the domain of VE-assisted training of spatial behavior to respond to this paper with critical comments and with their own thoughts about how best to advance knowledge and applications in this challenging domain.

4 ACKNOWLEDGMENT

This work was supported by ONR Grant N00014-96-1-0379. We are indebted to Terry Allard for his interest in, and contributions to, our work in this area.

5 REFERENCES

Aginsky, V., Harris, C., Rensink, R., and Beusmans, J. (1996). *Two Strategies for Learning a Route in a Driving Simulator.* Cambridge Basic Research Center, Nissan Research and Development, TR 96-6.

Allen, G. (1999a). Cognitive Abilities in the Service of Wayfinding: A Functional Approach. *Professional Geographer*, in press.

Allen, G.L. (1999b). Spatial abilities, Cognitive Maps, and Wayfinding: Bases for individual differences in spatial cognition and behavior, in R.G. Golledge (Ed.), *Wayfinding Behavior: Cognitive mapping and other spatial processes.* Johns Hopkins University Press.

Allen, G. L., Kirasic, K. C., Dobson, S. H., Long, R. G., & Beck, S. (1996). Predicting environmental learning from spatial abilities: An indirect route. *Intelligence*, 22, 327-355.

Allen, R. C., McDonald, D. P., & Singer, M. J., (1997). Landmark direction and distance estimation in large scale virtual environments. *Proceedings of the Human Factors and Ergonomics Society 41st Annual Meeting*, 1213 - 1217.

Arthur, E., Hancock, P. & Telke, S. (1996). Navigation in virtual environments. *Proceedings from the 10th Annual International AeroSense Symposium*, Florida.

Baenninger, M., and Newcombe, N. (1989). The role of experience in spatial test performance: a meta analysis. *Sex Poles*, 20(5/6), 327-344.

____________ (1995). Environmental input to the development of sex-related differences in spatial and mathematical ability. *Learning and Individual Differences*, 7(4), 363-379.

___________ (1996). Sauce for the goose, sauce for the gander. *Learning and Individual Differences*, 8(1), 65-68.

Bailey, J. H. (1994). Spatial knowledge acquisition in a virtual environment (Doctoral dissertation, University of Central Florida). *Dissertation Abstracts International*, 55(6-B), 2421.

Bailey, J.H., and Witmer, B.G. (1994). Learning and Transfer of Spatial Knowledge in a Virtual Environment. *Proceedings Human Factors and Ergonomics Society, 38th Annual Meeting*, 1158-1162.

Barfield, W., and Furness, T.H. (eds.). (1995) *Virtual Environments and Advanced Interface Design.* Oxford University Press.

Barfield, W., & Kim, Y. (1991). Effect of geometric parameters of perspective on judgements of spatial information. *Perceptual and Motor Skills*, 73, 619 - 623.

Bever, R. (1992). The logical and extrinsic sources of modularity. In M. Gunnar & M. Maratsos (Eds.), *Modularity and constraints in language and cognition*. Vol. 25 of Minnesota Symposium on Child Psychology. Hillsdale, NJ: Erlbaum.

*Bliss, J.P., Tidwell, P.D., and Guest, M.A. (1997). The Effectiveness of Virtual Reality for Administering Spatial Navigation Training to Firefighters. *Presence* 6(1), 73-86.

Boles, D. B. (1980). X-linkage of spatial ability: A critical review. *Child Development,* 51, 625 - 635.

Caplan, P.J., MacPherson, G.M., and Tobin, P. (1985). Do sex-related differences in spatial abilities exist? A multilevel critique with new data. *American Psychologist*, July, 786-799.

Chance, S.S., Gaunet, F., Beall, A.C., and Loomis, J.M. (1998). Locomotion Mode Affects the Updating of Objects Encountered During Travel: The Contribution on Vestibular and Proprioceptive Inputs to Path Integration. *Presence* 7(2), 168-178.

Chipman, S.F., Brush, L.R., and Wilson,D.M. (Eds)., *Women and Mathematics: Balancing the Equation.* Hillsdale, NJ: Erlbaum.

Christensen, R.R., Hollerbach, J.M., Xu, U., and Meek, S.G. (2000). Inertial Force Feedback for a Locomotion Interface, *Presence* 9(1).

Colle, H.A., and Reid, G.B. (1998). The Room Effect: Metric Spatial Knowledge of Local and Separated Regions. *Presence* 7(2), 116-128.

Connor, J.M., and Serbin, L.A. (1985). Visual-Spatial Skill: Is it important for mathematics? Can it be taught? in S.F. Chipman, L.R. Brush, and D.M. Wilson (Eds)., *Women and Mathematics: Balancing the Equation.* Hillsdale, NJ: Erlbaum.

Cronbach, L.J. *Essentials of Psychological Testing.* New York: Harper and Row.

*Darken, R.P. and Banker, W.P. (1998). Navigating in Natural Environments: A Virtual Environment Training Transfer Study. *Proceedings of VRAIS '98*, 12-19.

Darken, R.P. and Sibert, J.L. (1993). A Toolset for Navigation in Virtual Environments. *Proceedings of ACM User Interface Software and Technology*, 157-165.

Darken, R.P. and Sibert, J.L. (1996a). Navigating Large Virtual Spaces. *International Journal of Human-Computer Interaction*, 49-71.

Darken, R.P. and Sibert, J.L. (1996b). Wayfinding Strategies and Behaviors in Large Virtual Worlds, *Proceedings of Computer Human Interfaces Conference*, 1996 (*CHI'96*), 142-149.

Department of the Army (1969). *Survival, Evasion and Escape.* Field Manual #21-76. Washington, D.C.

____________ (1987). *Map Reading and Land Navigation.* Field Manual #21-26. Washington, D.C.

DiZio P, and Lackner, J.R. (1998) *Alleviation of motion sickness and postural instability during and after virtual environment exposure.* NAWCTSD Technical Report 98-009, Naval Air Warfare Center, Orlando, FL

Dorsey, J., and McMillan, L. (1998). Computer graphics and architecture: state of the art and outlook for the future. *Computer Graphics* 32(1): 45-48.

Durlach, N.I., and Mavor, A.S. (eds.). (1995)., *Virtual Reality: Scientific and Technological Challenges.* National Academy Press.

Fletcher, B. and Roberts, B. (1998) Training ROV Pilots with a VE-based Intelligent Tutoring System, *Proceedings of the 1998 Interservice/Industry Training, Simulation and*

Education Conference, Orlando, FL.

Gabbard, J.L., and Hix, D. (1997). *A Taxonomy of Usability Characteristics in Virtual Environments*, Final Report, ONR Contract N00014-96-1-0385, Department of Computer Science, Virginia Polytechnic Institute and State University, Blacksburg, VA [http://csgrad.cs.vt.edu/~jgabbard/ve/taxonomy]

Geary, D.C. (1995). Sexual Selection and Sex Differences in Spatial Cognition. *Learning and Individual Differences*, 7(4), 289-302.

Golledge, R.G. (ed.). (1999). *Wayfinding Behavior: Cognitive Mapping and Other Spatial Processes.* Johns Hopkins University Press.

Harris, L.J. (1978). Sex differences in spatial abilities: possible environmental, genetic, and neurological factors. In M. Kinsbourne (Ed.), *Asymmetrical Function of the Brain*, 405-522, New York: Cambridge University Press.

Henry, D.P. (1992). *Spatial Perception in Virtual Environments: Evaluating an Architectural Application.* MSEE thesis, University of Washington.

Imperato-McGinley, J., Pichardo, M., Gautier, T., Voyer, D., & Bryden, M. P. (1991). Cognitive abilities in androgen insensitive subjects: Comparison with control males and females from the same kindred. *Clinical Endocrinology*, 34, 341 - 347.

Just, M. A., & Carpenter, P.A. (1985). Cognitive coordinate systems: Accounts of mental rotation and individual differences in spatial ability. *Psychological Review*, 92, 137-172.

Kimura, D. (1996). Sex, sexual orientation and sex hormones influence human cognitive function. *Current Opinion in Neurobiology*, 6, 259-263.

Knott, J.A., and Gaze, C.E. (1996). Can spatial training erase the gender differences on the water-level task? *Psychology of Women Quarterly*, 20, 549-567.

*Koh, G., von Wiegand, T.E., Garnett, R.L., Durlach, N.I., and Shinn-Cunningham, B. (1999). Use of Virtual Environments for Acquiring Configurational Knowledge about Specific Real-World Spaces: I. Preliminary Experiment. *Presence* 8(6).

Landau, B. & Spelke, E.S. (1985). Spatial knowledge and its manifestations. In H. Wellman (Ed.), *Children's searching: The development of search skill and spatial representations*, 27-52. Hillsdale, NJ: Erlbaum.

Lewis, R.W. (1996). *Generating Three-Dimensional Building Models from Two-Dimensional Architectural Plans.* M.S. Thesis, Department of Electrical Engineering and Computer Science, Berkeley.

Linn, M.C., and Petersen, A.C. (1985). Emergence and Characterization of sex differences in spatial ability: A meta-analysis. *Child Development*, 56, 1479-1498.

Lohman, D. F. (1988). Spatial abilities as traits, processes, and knowledge. In R. Sternberg (Ed.), *Advances in the psychology of human intelligence* 4: 181-248. Hillsdale, NJ: Erlbaum.

Loomis, J.M., Golledge R.G., and Klatzky, R.L. (1998). Navigation System for the Blind: Auditory Display Modes and Guidance." *Presence* 7(2), 193-203.

Loomis, J.M., Klatzky, R.L., Golledge, R. G., Cincinellli, J.G., Pellegrino, J.W., & Fry, P. A. (1993). Nonvisual navigation by blind and sighted: Assessment of path integration. *Journal of Experimental Psychology: General*, 122, 73-91.

Lorenz, C. A., & Neisser, U. (1986). *Ecological and psychometric dimensions of spatial ability.* Emory Cognition Project Report No. 10, Emory University, Atlanta, GA.

Lowry, R., and Sidney, K. (1989). *Orienteering Skills and Strategies.* North York, Ontario: Orienteering Ontario.

Maccoby, E.E., and Jacklin, C.N. (1974). *The Psychology of Sex Differences.* Stanford: Stanford University Press.

Madden, S., and von Wiegand, T.E. (1999). TOADS: A two-dimensional open-ended architectural database system. *Presence*, submitted.

May, M., Peruch, P., and Savoyant, A. (1995). "Navigating in a Virtual Environment with Map-Acquired knowledge: Encoding and Alignment Effects. *Ecological Psychology* 7(1), 21-36.

McDonald, T. P., & Pellegrino, J. W. (1993). Psychological perspectives on spatial cognition. In T. Gärling, and R. G. Golledge (Eds.) *Behavior and Environment: Psychological and Geographical Approaches*, 47 - 82. N. Holland: Elsevier.

McGee, M. G. (1979). Human spatial abilities: Psychometric studies and environmental, genetic, hormonal, and neurological influences. *Psychological Bulletin, 86,* 889 - 918.

Mumaw, R. J., & Pellegrino, J. W. (1984). Individual differences in complex spatial processing.

Journal of Educational Psychology, 76, 920-939.

Neale, D. C. (1997). Factors influencing spatial awareness and orientation in desktop virtual environments. *Proceedings of the Human Factors and Ergonomics Society 41st Annual Meeting*, 1278 - 1282.

Paul, D. (1991). *The Green Beret's Compass Course*, 6th edition. Woodland, CA: Path Finder Publications.

Pausch, R., Proffitt, D., and Williams, G. (1997). Quantifying Immersion in Virtual Reality. *Computer Graphics Proceedings*, Annual Conference Series.

Peruch, P., Vercher, J.-L., and Gauthier, G.M. (1995). Acquisition of Spatial Knowledge through Visual Exploration of Simulated Environments. *Ecological Psychology* 7(1), 1-20.

Peterson, B. N., Wells, M., Furness, T. A., & Hunt, E. (1998). The effects of the interface on navigation in virtual environments. *Proceedings of the Human Factors and Ergonomics Society 42nd Annual Meeting*, 1496 - 1500. Santa Monica, CA: Human Factors and Ergonomics Society.

Pioch, N., Roberts, B., and Zeltzer, D. (1997) A Virtual Environment for Learning to Pilot Remotely Operated Vehicles, *Proceedings of the Virtual Systems and Multimedia '97 Conference.*

Plomin, R., and DeFries, J.C. (1998). The genetics of cognitive abilities and disabilities: investigations of specific cognitive skills can help clarify how genes shape the components of intellect. *Scientific American*, May, 40-47.

Rivers, C., and Barnett, R.C. (1999). *The Future of the Sexes*. Cambridge: Harvard University Press.

RLE Annual Report (1998). Sensory Communication Group, VSPAN Report.

Richardson, A.E., Montello, D.R., Hegarty, M. (in press). Spatial knowledge acquisition from maps and from navigation in real and virtual environments. *Memory and Cognition.*

Rossano, M. J., & Moak, J. (1998). Spatial representations acquired from computer models: Cognitive load, orientation specificity and the acquisition of survey knowledge. *British Journal of Psychology*, 89, 481 - 497.

Ruddle, R.A., Payne, S.J., and Jones, D.M. (1997). Navigating Buildings in 'Desk-top' Virtual Environments: Experimental Investigations Using Extended Navigational Experience. *Journal of Experimental Psychology: Applied*, 3(2), 143-159.

Ruddle, R.A., Payne, S.J. and Jones, D.M. (1998). Navigating Large-Scale 'Desk-Top' Virtual Buildings: Effects of Orientation Aids and Familiarity. *Presence* 7(2), 179-192.

Ruddle, R.A., Randall, S.J., Payne, S.J., and Jones, D.M. (1996). Navigation and Spatial Knowledge Acquisition in Large-Scale Virtual Buildings: An Experimental Comparison of Immersive and Desk-Top Displays. *Proceedings of the Second International FIVE Conference*, 125-136.

Satalich, G.A. (1995). *Navigation and Wayfinding in Virtual Reality: Finding the Proper Tools and Cues to Enhance Navigational Awareness.* MSEE thesis, University of Washington.

Seidel, J.J., and Chatelier, P.R. (eds.). (1997). *Virtual Reality, Training's Future? Perspectives on Virtual Reality and Related Emerging Technologies.* Plenum Press.

Silverman, I. & Eals, M. (1992). Sex differences in spatial abilities: Evolutionary theory and data. in J. Barkow, L. Cosmides, & J. Tooby (Eds.), *The Adapted Mind: Evolutionary psychology and the generation of culture* (pp. 533-549).

Stanney, K.M. (ed.). (1999). *Handbook of Virtual Environments.* Lawrence Erlbaum and Associates.

Stiles, R. (1999). *Virtual Environments for Training*, Final Report, ONR Contract N00014-95-6-0179, Advanced Technology Center, Lockheed Martin Missiles and Space, Palo Alto, CA.

Stoakley, R., Conway, M.J., and Pausch, R. (1995). Virtual Reality on a WIM: Interactive Worlds in Miniature. *CHI'95 Mosaic of Creativity.*

*Tate, D.L., Sibert, L., and King, T. (1997). Virtual Environments for Shipboard Firefighting Training. *Proceedings of the VRAIS*, Albuquerque, N.M., 61-68.

Templeman, J. , Denbrook, P.S., and Sibert, L.E. (1999). Virtual Locomotion: Walking-in-Place through Virtual Environments. *Presence* 8(6).

Tlauka, M. and Wilson, P.N. (1996). Orientation-Free Representations from Navigation through a computer-simulated Environment. *Environment and Behavior*, 28(5), 647-664.

Waller, D.A. (1999). *An assessment of individual differences in spatial knowledge of real and virtual environments.* Unpublished dissertation, University of Washington.

*Waller, D., Hunt, E., and Knapp, D. (1998). The Transfer of Spatial Knowledge in Virtual Environment Training. *Presence* 7(2), 129-143.

Wilson, P. N., Forman, N., Gillett, R., & Stanton, D. (1997). Active versus passive processing of spatial information in a computer-simulated environment. *Ecological Psychology,* 9(3), 207 - 222.

Wilson, P. N., Foreman, N., & Tlauka, M. (1997). Transfer of spatial information from a virtual to a real environment. *Human Factors,* 39(4), 526 - 531.

*Witmer, B.G., Bailey, J.H., and Knerr, B.W. (1995). *Training Dismounted Soldiers in Virtual Environments: Route Learning and Transfer.* USARI Technical Report 1022.

*Witmer, B.G., Bailey, J.H., Knerr, B.W., and Parsons, K.C. (1996). Virtual spaces and real world places: transfer of route knowledge. *International Journal of Human-Computer Studies* 45, 413-428.

Witmer, B.G. and Kline, P.B. (1998). Judging Perceived and Traversed Distance in Virtual Environments. *Presence* 7(2), 144-167.

http://www/cs/berkeley.edu:80/~rickl/BMG/

http://www.lcs.mit.edu/research/projects/project?name=9919

http://www.lcs.mit.edu/research/projects/project?name=9946

http://graphics.lcs.mit.edu/index.html

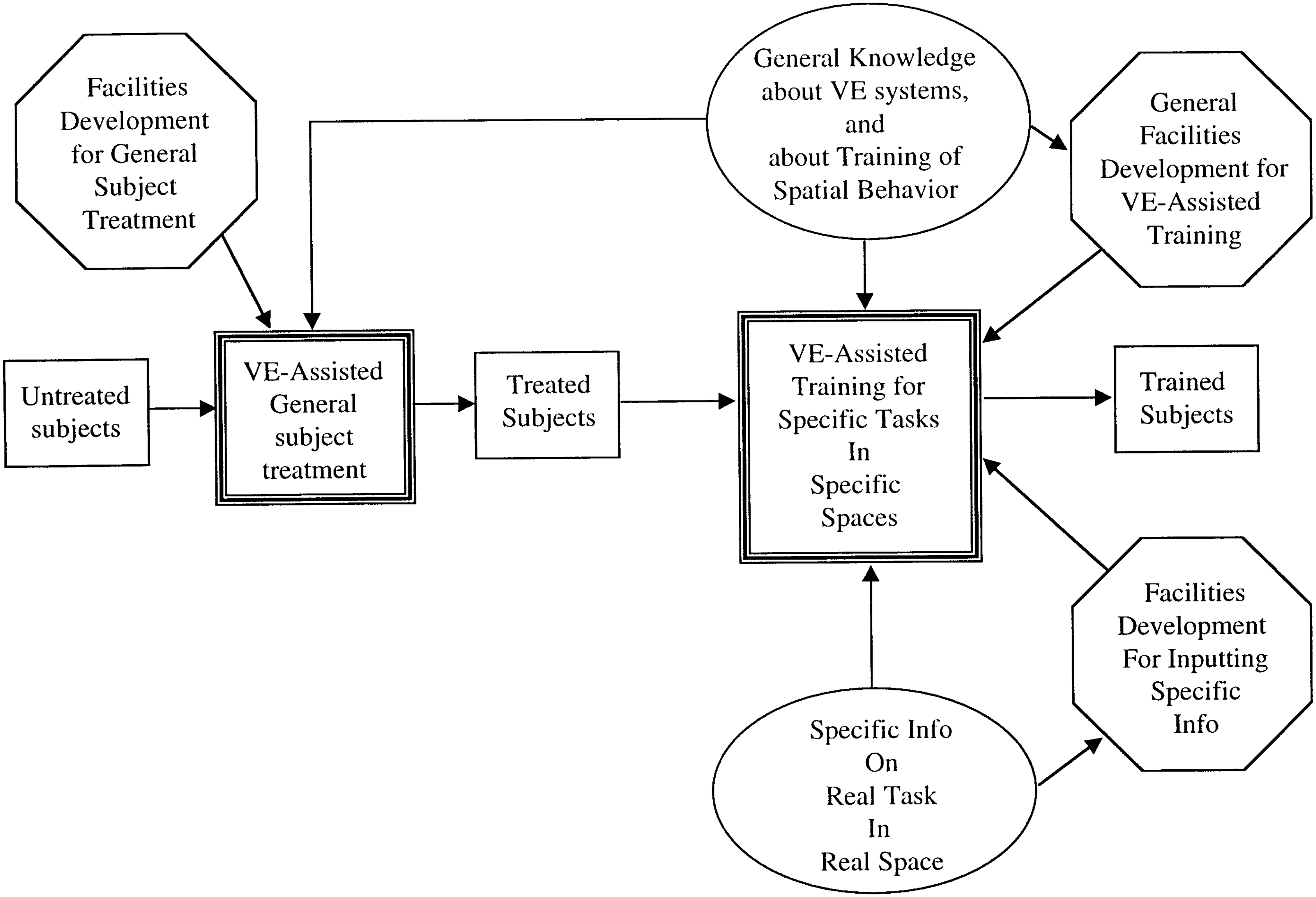

Figure 1. Schematic of Program Elements

page 41

A RANDOM WALK: Mastering Musketaquid

Rebecca Lee Garnett

Research Laboratory of Electronics
Massachusetts Institute of Technology

During Year 2 of the ONR-funded VSPAN project (1997), one of the team members who has a personal interest in the cognitive processing of spatial information began a series of self-directed terrain exploration experiments, which were intended to provide direct experience of, as well as documentation of, the developmental stages an individual may encounter in the process of acquiring knowledge about a specific large-scale real-world space. The area selected for this series of experiments, which has continued through the current year of the project, is a 36-square mile tract of land in Middlesex County, Massachusetts, centering on the townships of Concord and Lincoln, and including portions of the contiguous towns of Acton, Carlisle, Bedford, Weston, Wayland, and Sudbury.

A number of advantages recommended this area as a prospective venue for orientation and wayfinding exercises. First, it is accessible by several means of public transportation. Second, it offers within a comparatively small area a high proportion of undeveloped land, maintained and protected in both the public and private sectors by local, state and national conservation agencies, educational institutions and land trusts. Third, while its terrain is not excessively challenging physically, its glacial past has given it a rugged and complex array of land conformations and habitats, including more than a dozen drumlins and bedrock outcroppings greater than 200 feet in altitude, three rivers, a number of lakes, wetlands and brook watersheds bounded by steep rocky eskers, well over one hundred miles of trails and footpaths, and finally, extensive forest tracts, most notably the Adams-Walden-Flint's Pond woods in Lincoln, and Concord's 650-acre Esterbrook Woods, which serves as an ecological research station for Harvard University.

Man-made streets and roads in the vicinity are for the most part safe for pedestrian travel, and offer a unique challenge to the city-block spatial thinker in that they are laid out in a triangular rather than a square grid pattern, converging at hubs or "Y" intersections rather than at conventional right-angle corners.[1] Woodpaths offer a similar spatial challenge, as many of them are circular or sinusoidal in their layout, following natural contour lines of the terrain as first etched into the ground by the animals that once ranged freely in this historic Native American hunting ground.

Much of the area is well documented in locally-produced trail guides and maps, which can be used in conjunction with commercially available street maps and USGS

[1] See, for an analysis of modern traffic problems related to this layout, "Traffic woes have roots i n the past," Concord Journal, March 12, 1998.

terrain maps to provide the increased level of detail which is helpful to a beginner in the process of map-based route learning or wayfinding. In addition, there is an abundance of textual, or word-based, information describing the territory and its network of roads, trails and footpaths (many of which, due primarily to the preservation efforts of the respective towns, have not appreciably altered in the last 150 years), in the writings of the New England Transcendentalists, including
Emerson, Hawthorne, Channing, Alcott, and Thoreau.

Walking May Be a Science . . .[2]

Between May 1997 and December 1999, more than one hundred full and half-day trips were taken to the experimental venue, which for purposes of the project was called by the experimenter the "spatial laboratory." Historically, the six-mile square was the unit of land sold by the Native American people of the Musketaquid River to the European settlers who established, at the confluence of the river's north and south branches, the colonial village of Concord. The experimenter had no prior exposure to, or knowledge of, the landscape or terrain features of the venue.

Some free exploration without the use of maps was undertaken at the beginning of the project, centering around the approximately one square mile of Concord Village. Beginning in November, 1997, a series of eight to fourteen-mile hikes were planned and executed, extending through and around the limits of the territory, with each route planned on a terrain map and studied in advance. One set of routes followed improved roads, and focused on learning man-made locations, landmarks and boundaries in the area; a second set focused on trail hiking and the process involved in learning unimproved footpaths and natural landmarks, terrain patterns, and ecological habitats. Trips took place at approximate one-week intervals. All exploration was done on foot, without the potentially distracting company of a companion. No compensation was received for time spent on the walks, and transportation and materials costs were not reimbursed for any portion of these experiments.

A USGS topological map covering the appropriate quadrant(s) was carried on all trips, but was consulted only at decision points during the trip, during terrain visualization exercises (see below), or to identify an unknown object in the line of sight.[3] In the spring of 1999, an engineer's compass was acquired and utilized with the maps on a number of walks in wooded terrain. Binoculars were also carried and used to bring into focus landmarks in the distance or on the horizon. A street map was consulted after each trip, and the path taken on each day was highlighted with a yellow marker. The length of the route was determined after each trip by running a

[2]Section headings are taken from Thoreau's writings; the full quotations and the sources are given at the end of this document.

[3]The entire area explored can be found on the Maynard and Billerica quadrant maps of the U.S. Geological Survey.

mechanical distance calculator along the route drawn on the street map. Although formal time checks were not made along any route, the full-day hikes did not exceed 7.5 hours, and tended to range between 5 and 6 hours in duration.

Fig. 1. The confluence of the Assabet and Sudbury Rivers in Concord., Massachusetts. Lowell Road is seen as a diagonal curved line crossing the river. near the center of the image. Satellite photo courtesy MIT/MassGIS Orthophoto Server.

After November 1997, field notes of observations were made on all walks, describing such features as ground conditions along the path, major route decision points, lines of sight where applicable, significant landmarks, both man-made and natural, and any physical sensations (non-visual cues) contributing to the experimenter's understanding of place (i.e., a suddenly cold local wind in the woods, which could indicate the presence of a large tract of ice - a frozen river - in the vicinity; the sound of frogs singing, which could indicate the location of a nearby wetland hidden from sight below the path; road noise, indicating the direction and proximity of a highway; the scent of a flowering tree, which could serve differentiate one trail head from another when such openings are in close proximity and visually similar, and the like).

No walk was an exact duplicate of any other walk previously made, although the same regions might be visited in somewhat similar order, and except in a few notable instances no part of the route was retraced in a single walk. Care was taken to follow major roads and paths in opposite directions at one or more different times, and a number of localities explored originally in summer were revisited in winter, when the leaflessness of trees and absence of other obscuring vegetation allows for a less unobstructed line of sight to nearby landmarks, and when snow makes the inequalities of the terrain more immediately visible (i.e., a nearby hill, invisible through summer foliage, becomes more obvious in winter, when a snow cover lies on its upward sloping face). Although written notes were kept on each trip, they were not consulted in preparing for any subsequent trips to the same general area, and served primarily as a tool for keeping the experimenter's attention focused on observation and visual image construction while on the path.

The Discipline of Looking ...

The purpose of these experiments was twofold: first, to learn and apply the technical competencies involved in map reading, spatial orientation and wayfinding, as set forth in a number of training manuals and outdoor-oriented publications; and second, in so far as possible, to experience and document one individual's personal development of configurational knowledge of a specific large-scale natural space. It was hoped that this personal exploration would give the experimenter an immediate insight into the nature of spatial learning, and thereby assist in the design of further experiments and formal training programs in which the principles of spatial knowledge acquisition could be taught in a clear, direct, and cost-effective manner.

The experimenter had no prior training in map interpretation, wayfinding or route knowledge acquisition, and indeed is one who had tested poorly in school on paper-based standardized measures of spatial orientation and the manipulation of 2D object representations in 2D space. By self-report, the experimenter was not what might be termed "high-visual," tending to avoid complex visual stimuli in daily life, expressing a preference for verbal directions over graphical maps, and admitting to difficulty with creating and manipulating mental imagery. This project, then, was ultimately a personal attempt to learn whether or not a large, complex natural space of this kind could actually be "mastered" in the face of such challenging limitations, and if so, whether the insights gained in the process would prove useful in organizing a program designed for teaching others, not simply the configuration of a specific space or set of spaces, but general principles of learning about spaces as well.

Any self-directed and consciously self-altering process can yield material which is only anecdotal at best, but here follow brief summaries of some techniques tried and insights gained through the course of this project, many of which arose out of the valuable experience of making "mistakes" and getting lost along the way. Because the technical discipline of wayfinding has been amply set forth in a number of publications, this discussion will focus on what the experimenter found to be personally meaningful in learning a specific space, and, more generally, in developing the qualities of a "spatial mind."

Attention. Much of the basic cognitive work done along the paths seemed to have less to do with the specifics of "route learning" as such than with the development of continuous, route-associated situational awareness. This process seemed to have a great deal in common with the mental discipline involved in the meditative practice of mindfulness: focusing the attention and maintaining a sense of "presence" at all times to the full sensory quality of the environment as one travels through it. Letting one's mind wander for even a minute or two could have the immediate result of "getting lost" - or rather "coming to one's senses" without the faintest idea where one was or how one got there. In a forest maze of forking pathways and semicircular loops with branching termini, this is a matter of no small seriousness, even if one knows the woodland is less than a hundred acres in extent. Disciplined attention provides the means to keep a path history in mind, allowing for the retracing of steps, if necessary, to a recognizable point from which the route can be resumed. And since even the best map will not necessarily represent all the pathways, and therefore all the decision points, to be encountered along the ground, continual conscious focus may provide the only compensation for missing information or for confusing matters of scale provided by the map.

Observation. A second level of learning involves training the eye to notice and interpret inconspicuous objects or patterns of objects in the environment: on the ground along the pathside, above one's sight-line, on the horizon. In natural space, these small landmarks might well take the form of individual plants, terrain features such as rock outcroppings, or the evidence of an animal's activity, a hole in a bank, a nest in a tree, or the like. The process of watching what the eye selects to use as a landmark in passing, or, conversely, how a potential landmark recommends itself to consciousness amongst all the objects available to view, involves a concentration not less intense and continuous than that of training the attention to maintain a conscious awareness of one's path history at all times. And in turn, each landmark can then serve as a constellating object around which previously undifferentiated space can be organized, helping to form an orienting network of small territories, or districts, which overlays, supplements and amplifies the "route" established on the paper map and followed subsequently along the ground. Training the eye to take lines of sight at crossroads and other decision points, to be aware of the length and direction of shadows, to follow the lines of man-made landscape features, such as stone walls, and the natural contours of brooks, ridges and hillsides, allows for the connection of isolated, static landmark information into a synthetic, more nearly three-dimensional model of the space, and serves as a transitional stage between a linear apprehension of the territory (the route) and a full, comprehensive, configurational appreciation of it (the survey) - mediating the cognitive shift from egocentric to exocentric understanding.

Memory. A third component to this work, and one which may potentially prove the most controversial, involves the development of a strong and flexible visuo-spatial memory. Such a process might begin with the conscious act of apprehending an object, taking a visual impression of it with such clarity, fidelity, and detail that it can be recalled to mind afterwards and held in its appropriate place in a sequence of other such intentional and clear visual impressions. Whether it be a single landmark, a circumscribed vista, or a broad and open landscape viewed from a hilltop, the impression so registered in memory can become an anchoring point for the imaginal creation of a fully three-dimensional mental representation of the territory. Helpful in developing this faculty is the process of taking notes on objects observed and designated as landmarks: the concentration involved in noting one's observations serves as an aid to the establishment of a vivid memory, and the review of notes afterwards as a potential aid to recall.

Take a Dozen Steps . . .

Thus far the discussion has focused on the mental skills which are applicable to the process of learning about space in a general sense, and which are equally applicable to achieving a clear sense of "presence" anywhere in the course of one's daily life. Focused attention, the ability to observe, and a strong, flexible memory are cognitive attributes which are not limited to specific learning situations, but can certainly be included as components in the development of what might be called "spatial consciousness" or "spatial mind." The present project, while demonstrating that these three cognitive attributes are indeed amenable to improvement through conscious practice, also showed that they can be developed and strengthened by simple, cost-free, and essentially non-technical mental practices.

But additional work beyond the development of these general mental attributes was necessary to produce the consolidation of what were essentially linear impressions, encountered sequentially along a number of learned routes, into a comprehensive, integrated, seamless and unitary configurational model of the territory under study. Following are some brief descriptions of exercises the experimenter used and found helpful in accomplishing this cognitive shift from route to survey knowledge of the space - transforming the centrality and mobility of the observational viewpoint to a static, comprehensive, synthetic sense of the whole environment.

Once a number of paths have been laid down in memory, the terrain between and around them can be consolidated mentally through the exercise of seeking alternate routes - a practice which can be followed on the ground as well as after the conclusion of a walk. Asking oneself questions along the way can guide this process, the act of questioning providing for a continuous, conscious engagement with the circumstances of the unfolding route. "If this path were flooded, or overgrown with vegetation, how else could I get to my chosen destination?" "If this bridge were damaged, how might I get over this river?" That these are not idle considerations is shown by the fact that during the period of this experiment, one of the historic bridges carrying a state highway across the Sudbury River between Concord and Lincoln did collapse, a situation which necessitated the rerouting of traffic for four days while a temporary structure was put in place. Seasonal flooding makes many of the riverside and meadow paths in the area inaccessible; and it would be only a small exaggeration to say that throughout the territory, the shortest distance between any two points lies through a swamp, which one may or may not be appropriately outfitted to traverse.

More complex than the process of seeking alternative routes, though in some sense related to it, is the process of imagining a specific location from a different point of view; asking for example, "how might this hill look if I approached it from the other side? What might I see if I got to the top of it?" "Where do I think this winding path will intersect the road?" "How would this street look if I were standing at the other end of it?" Although this technique may be the same as that which in theory is measured on the paper test of mental rotation or translation, it has the advantage, for the learner, of operating within an environment rich in informational detail, providing more intrinsic interest than empty abstract line drawings of objects presented on paper. And such mental images, once formed and transformed, can be easily verified in the actual environment, assuming that no insuperable barrier intervenes.

Related to the concepts of visual memory development and alternate route-seeking is the technique of visualizing, in advance, the likely habitat represented by legend coding on a terrain map. A smooth green block bisected by a wavy blue line on the map might well prove to be a virtually impenetrable alder thicket eight feet high and soggy underfoot when one reaches it on the ground, impassable even though a path - on the map - may run through it. Conversely, a prominent blue line representing a stream on the map may be quite invisible from the path, if it is one that dries up seasonally, disappears underground, or has been piped through a conduit for that portion of its length. For those who have difficulty, as the experimenter did, in visualizing a situation when given an abstract representation of it, some amount of prior experience in seeing the terrain represented can be helpful in learning to develop the imaginal capacity needed to construct, under similar circumstances, a reasonably accurate mental image of a space.[4]

A further exercise helpful in establishing a survey sense of the territory involves sitting with a terrain map in a designated spot on a path, preferably a closed-in spot in the woods, laying out the map so that it is oriented to conform with the view that is directly ahead, and tracing out a line straight ahead on the map while trying to visualize what one might see, if the vista from where one sits were unobstructed. This can be done either with or without reference to compass direction (i.e. - "what is directly north of me as I sit here?", or, "what is directly in front of me here?"). How would I get to it, assuming that the most direct way - straight ahead - is not actually

[4]It is surprising how much work is involved in getting beyond the documentation and the expectations that may be pre-set by the study of documentary material. The importance of this cannot be overstressed. The internalization of an inadequate map may give worse results on the ground than no prior preparatio at all. Furthermore, there is a need during the preparatory or study phase to think in terms of senses other than the visual. The author, in preparing to do field work in Brazil some years ago, viewed many films and videotapes prior to departure - but none of them provided adequate preparation for the influx of smells and air texture information encountered on the ground. The suddenness and primacy of these sensory inputs can quite eclipse visual impressions and visually-based memory, something to keep in mind when image-oriented training is administered.

unobstructed or available? An alternative form of this exercise involves stopping at random on the path and pointing, as best one can, in the direction of some central or well-known feature of the landscape. In the spatial laboratory, for several weeks this took the form of asking "Where's Walden?" from various spatially-distributed locations in the area. This imaginal technique proved to be of great help in the process of linking the many small, discrete regions of the space together into larger, composite districts, in a way that travelling directly from one to the other (which may not have been physically possible) might not so easily have done.

In order to comprehend the network of streets and roads traversing the space as it actually presents itself to the user - pedestrian or driver - it proved to be of vital importance to abandon most if not all of the assumptions one habitually makes about street systems, derived from a lifetime's experience of getting around in a grid-based, city-block-style environment. The towns of Concord and Lincoln, as mentioned above, are laid out not in square blocks but in triangles, with major roads radiating out from the town centers like spokes from a giant wheel. As each road progresses across the territory at its angle, widening its distance from each of the others, it forks into subsidiary roads, until the pattern created resembles that made by arteries and veins in the body, or the branching of rivers and streams, more than the customary straightforward layout of urban streets and suburban roads.

Furthermore, throughout the entire territory, because of the peculiar history of highway construction from colonial times to the present, a road which continues straight through a fork cannot be assumed to be the same road - more than likely, the original road itself has curved aside to create the apparent fork, and the straight-ahead road is a different one altogether, constructed at some later time. (To some degree, this is the case with many of the woodland trails as well, which at one time were cart-tracks or foot paths leading to farms or houses, no longer extant, in the woods.)

There are very few cross-roads intersecting these great angled thoroughfares, and a surprising dearth of signposts naming the intersecting ways, so mistaking one's way on the roads can lead to a very lengthy and tiring recovery, in some ways more frustrating than the experience of losing one's way in the woods. In this regard, a close attention to scale on the map, whether a street or terrain map, became more important than previously expected. The kind of route rehearsal which might take the simple form of studying a map and then narrating to oneself, while in the territory, in what direction to proceed, where to turn ('take the second path to the left"), and so on, can prove quite treacherous when some important detail is missing - when an expected turnoff or throughway is found not to exist, for example, or when a new path has been inserted where none was shown on the map. New housing construction occasionally posed unanticipated problems in finishing a walk along the route as planned, and some wood trails once open to the public were found to be restricted by the order of the individual landowners since the publication of the trail guides.

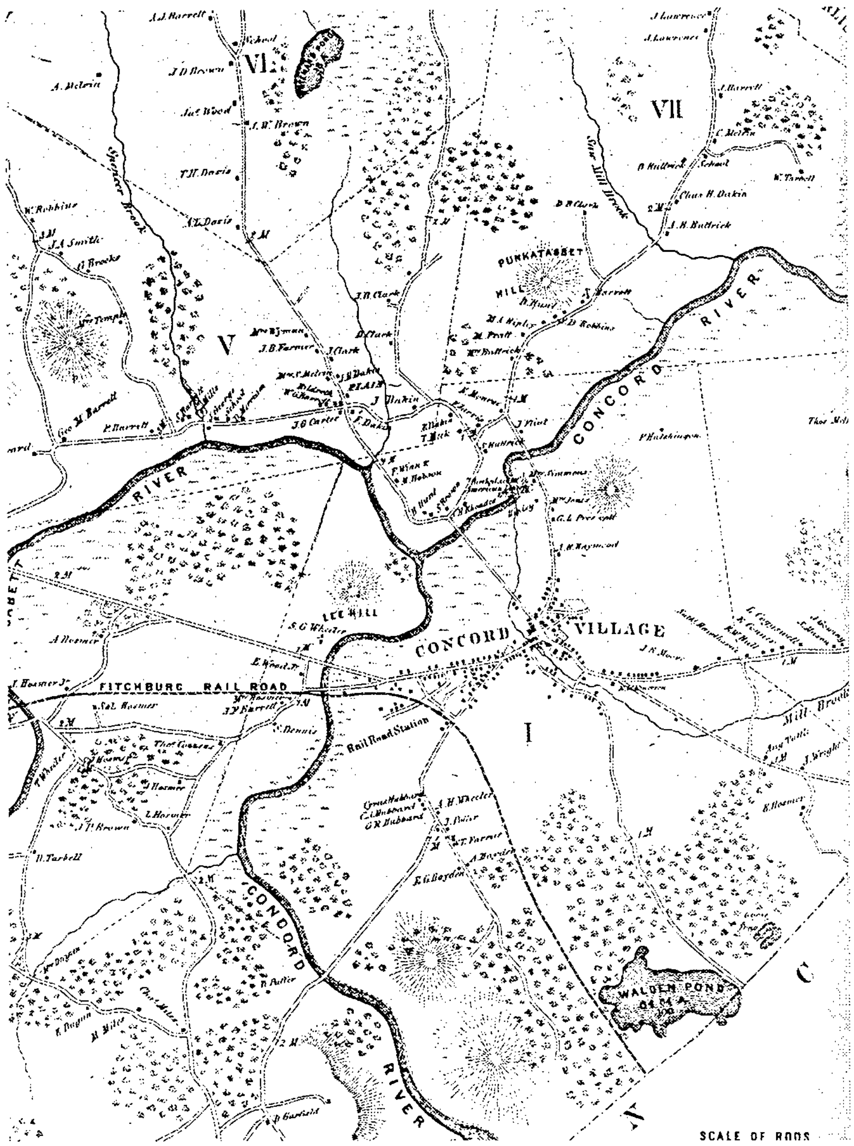

Fig. 2. A segment of the 1852 Walling map of Concord Township, showing the town's distinctive pattern of branching arteries and "Y"-shaped intersections. The area shown in Figure 1 is visible near the center of this map, which is reproduced in *A Thoreau Gazetteer*, by Robert Stowell and William Howarth.

As each discrete segment of the territory became more and more familiar, through physical experience as well as repeated visualization exercises, it became possible to characterize each one in terms of some unique aspect of its physical layout, the unique lesson that it had to teach about the process of spatial knowledge acquisition. Some of these lessons became clear in flashes of insight experienced on the spot; others came to consciousness more slowly and only after the path had been left for the day, and of these, some came only as the result of much puzzling over wrong turns, missed trailheads, or hopeless and complete misorientation. Foremost among these "teaching locales" are the Esterbrooks Country, known locally as a place where one can still get lost, no matter how many times one goes there; other important insight spots that might be singled out for mention are Smith's Hill in Lincoln (how does one know when one has gone halfway round a circle?); Concord's Conantum district (a desperate case of interlocking S-curves); the Flint's Pond/Saw Mill Brook tract in Lincoln (why does this path keep leading to somebody's back yard?), Bateman's Pond woods (identical bifurcated trail exits leading in opposite directions); and the Fair Haven woods, along the Concord-Lincoln border (why are these trails always so much longer than I remember?).

Configurational knowledge of a specific space - the comprehensive, complete and fully-dimensional mental model of the territory - while perhaps difficult to define in a theoretical sense, can be demonstrated at some level by a variety of measures, some of which might be quite ordinary or commonplace. In the current year of the project, after having walked perhaps 80% of the improved roads in the territory and 65-70% of the public woodpaths (which in the town of Lincoln alone stretch to more than 70 miles), the experimenter developed a number of small tests of the degree to which configurational knowledge of the space had been developed. These included: the ability to give directions, when stopped by lost motorists anywhere within the territory; a weekly review of the Concord Police and Fire Log, to identify and visualize the location of all sites mentioned in each report; and, in conjunction with the study described below, a regular, several-times-weekly visualization exercise based on routes through the territory as described 150 years ago in the journals of Concord naturalist Henry David Thoreau (see next section). To date, there has been no attempt to draw a map of the territory or any region within it, although this would be appropriate in the next stage of the project.

Finally, a number of "training transfer" explorations were undertaken in other localities, with good results, and several longer train excursions were taken in order to view significant features of the landscape from various locations outside the actual venue. A list of these trips is provided at the end of this report, together with a list of the routes and regions explored throughout the learning phase of the experiment.

A Greater Miracle . . .

It would be quite misleading to claim that the process of learning the configuration of this specific space was accomplished in an easy, quick, or truly cost-effective manner. Perhaps because of the grand scale of the region (36 square miles of hill, river, village, farm, woodland, swamp, meadow, and highway), the decision to explore the entirety of it on foot, and the experimenter's personal cognitive challenges and naivete with respect to the technical disciplines of spatial orientation and wayfinding, the process has taken the better part of two years - and there still remain some spaces on the map which have yet to be explored for the first time. Nevertheless, the experimenter has certainly achieved the "feel" of having mastered, to some degree, the terrain of the chosen territory, and, in addition, now has the confidence of being able to apply the aptitudes developed in this experiment to other locations and in other situations.

The approximately 500 hours spent in hiking on the ground and experiencing in person the paths and conformations of this complex environment were only a small portion of the actual time invested during these two years in the larger learning process, which continued through many additional hours of visualization, map study, route rehearsal and memory exercises. While the experimenter can be personally pleased with the results of all this labor, it is hardly to be recommended as a model for the rapid, cost-effective acquisition of knowledge about a large-scale natural space. And although basic principles and techniques of spatial learning can perhaps be articulated in a clear and concise form, there may be no viable substitute, in the case of a beginner to this practice, for a certain amount of time spent on the ground, learning by trial and error how to explore a space effectively - how to look, as well as what to look for, while exploring.

Once those techniques have been adequately established, though, the process of learning "specific" spaces could conceivably accelerated through the use of sequenced images, keyed in to 2- and 3D maps with various levels of detail, virtual (or material) environmental models, and even textual information of greater or less completeness. It is conceivable that in some real-life situations for which learning a specific space may be necessary, the only available information about that space may be presented in verbal form - the testimony of witnesses, the remembrance of former residents or visitors to the space, historical records, a set of directions which may or may not be complete or accurate. When moving from the (presumably) unambiguous graphical environment of a map to the semantic realm of narrated spatial information, one encounters a complex set of interpretational problems: two witnesses will not describe a given situation in the same way; perceptions of scale may vary; landmarks and even streets may have alternative names; key spatial elements may be remembered out of sequence. In daily life, very few individuals have the experience required to give a clear and concise set of directions even under the best conditions, and what one may choose to describe as a prominent reference feature in the landscape may be totally overlooked when the identical space is described by another.

In conjunction with the wayfinding work described in the previous section, the experimenter undertook a close reading of the published journals of Henry David Thoreau, whose exceptional knowledge of the landscape around his native town of Concord has been characterized by his British biographer, Henry Salt, as "an instinct for locality." Known almost exclusively for his two-year experiment in self-sufficient living at Walden Pond, Thoreau's subsequent work as a field ecologist has only recently begun to attract the scholarly attention it has long deserved, and among the more than 6,000 published pages of notes that he kept from his daily walks through the Concord countryside during the last twelve years of his life is evidence of an as-yet unrecognized, but quite profound, interest in what today we might recognize as the cognitive process underlying terrestrial navigation, situational awareness, and spatial knowledge acquisition.

For the purposes of this study, the experimenter approached the journals of Thoreau as a set of verbal descriptions covering many of the same routes, landmarks, lines of sight, ecological habitats and land conformations still available to be seen by anyone who walks in the vicinity. His meticulously detailed descriptions of the paths he took each day (naming the destination and a number of intermediate landmarks along the way), and the sights and sounds he encountered as he walked the ground of his territory, supplied the experimenter with a text-based accompaniment to the maps and trail guides studied in order to learn the space. Specific passages from the journals could serve, once an area had actually been explored by the experimenter, as an aid both to scene recall and to self-directed visualization exercises, in which the scene was consciously manipulated to conform to its probable appearance in Thoreau's lifetime. Although much of the terrain in the area is now protected with conservation restrictions, and the basic road network remains unaltered from his time, a great deal of change and development has taken place since then; construction has obliterated many of his favorite habitats, and a state highway, bisecting the territory with a great rushing arc of traffic, makes many of his favorite routes impossible to duplicate at the present time. Nonetheless, it was a surprise to the experimenter how many of Thoreau's walks can still be generally approximated, and how carefully he had noted the significant features of the terrain and its many vistas, consciously putting into practice the techniques of observation that he had taught himself through studying the writings of explorers and landscape painters, and conversing when he could with local hunters and trappers, timber prospectors, and Native Americans.

Because a great amount of his time was devoted to achieving an understanding of the life-cycle and development of individual plants, many of which he discovered by accident in out-of-the-way places on his walks or in the course of his professional activity as a surveyor, it was essential for him to develop a clear, though highly personal, system of geographic information, so that he could find the plant again on subsequent visits to its vicinity, or could, if he chose, direct others in such a way that they might also find it. Given the natural environment in which he did this work, the landmarks he typically chose to help him identify and remember the location of a plant of particular interest were often other notable plants in the vicinity; and in turn, the plant of interest often became in itself a landmark and a regularly named route destination, or lent its name to the locality or habitat in which Thoreau had found it.

In reading through twelve years of his journal notes, one finds Thoreau engaged in an almost continuous naming process, subdividing and segmenting regions of his territory with ever more appropriate and specific names, and at the same time districting, or combining his small individually-named regions into broader or more inclusive locales and environments. Some of his locational nomenclature persists in the area to this day - "Conantum," for example, was named by him, as was the Esterbrook Country. But for the most part his names for places in his countryside remained private, idiosyncratic, linked to the existence of a living plant or animal under his observation, and somewhat obscured by his habit of switching between common and scientific names to designate the same individual (for example, the destination indicated by the words "to lygodium" is the same as that noted as "to climbing fern;" "to epigaea" is the same as "to mayflower").

Not all of his walks were reported in equal detail, although the best of them can be traced along a topographical map with absolute fidelity.[5] So clear are some of his passages, in fact, that one could conceivably take a page of Thoreau's journal, and walk as though in his footsteps to the very site of the tree or bird's nest or fern-covered rock that he had made a special trip to observe, or look out, as through his eyes, from a hilltop along one of his favorite routes to the very same horizon he had so faithfully set down in words on that journal page. But during his long periods of illness, or when he was under the pressure of other obligations, his write-up of his field notes could be quite abbreviated, so cryptic, in fact, as to become incomprehensible for the reader who has no easy means of recalling all the information previously recorded about that route, or that specific destination, or who may not be familiar with the alternative naming conventions by which a single location can be referred to under a number of different and apparently unrelated names.

In order to resolve some of the ambiguity inherent in Thoreau's nomenclatural system, the experimenter developed a simple computer matrix for the manipulation of text-based geographic information in Thoreau's journal. The process of this component of the project is described in the section which follows.

With the Utmost Industry . . .

In January 1998, after the experimenter had attained to several months' familiarity with the actual space described in Thoreau's field notes, a searchable database was created, using Microsoft Excel, into which verbal information pertaining to the geography of the region could be entered, sorted, and manipulated along a number of predetermined dimensions. Because so much of Thoreau's landmarking practice utilized the names of individual plants or specific botanical habitats, the core of the database was built around the *Botanical Index to the Journals of Henry David Thoreau*, which was prepared by Harvard University's Ray Angelo and is available both in print and on the internet. Professor Angelo's determination of the scientific, species name of every plant mentioned in the journal was seen as the key to resolving the bewildering variety of common names used by Thoreau in his daily field work. The *lygodium palmatum*, already mentioned, could appear under the name "climbing fern"; the *equisetum hyemale* could be referred to as "scouring rush"; other plants might have three or more common names; yet each of these groups of names might signify one single location, perhaps even one individual intended to serve as a landmark, and thus for the purposes of accurate geographic distribution needed to appear as one coherent data set.

5See, for example, the route Thoreau and his friend W. E. Channing followed on their walks through the Inches Woods in Boxboro, traced by Alan B. Rowher on the website of the Boxboro Historical Society: http://www.multimgmt.com/BHSociety/Thoreau%20visits.htm.

Professor Angelo's index was amplified to include the names of agricultural plants appearing in the pages of Thoreau's work; these references had been omitted as being of no particular interest to botanists, but since many of them were identified with a specific plot of ground (e.g., "Wheeler's rye field"; "where Potter planted corn last year") ,they proved to be important elements in the organization of information related to plant-identified landmarks or locations in the vicinity under study.

In the early stages of his field work, Thoreau was in many cases as explicit as possible in setting down the particulars of his itinerary: his regular walk to Walden, for instance, might be taken "down the railroad causeway" or "via poorhouse road," or with a preliminary detour around the Fair Haven woods, or by way of Brister's Hill or Goose Pond. Once the botanical information had been laid down in the database as an anchor, this detailed itinerary or route information was then entered, in a separate column, for each of the days on which Thoreau's walking notes mentioned a vascular plant included in the database. Another column was then created to record the actual location of the plant named.

In instances in which Thoreau himself noted the plant's specific location on the day he observed it (for example, "at the top of Heywood's Peak at Walden"), this information was entered in a hierarchical form ("Walden: Heywood's Peak: top"). Where no such specific information was given, the plant location was entered as "no location." Whenever it was possible, through context, to extrapolate a probable location or region in which the plant was observed, this conjectural information was entered, with a question mark indicating the ambiguity involved ("between poorhouse and Goose Pond path?"). When the location of a specific plant, otherwise unrecorded, could be inferred from previous passages describing it, the conjectural location was indicated with a carat, to signify that information missing from the journal pages was filled in by the compiler of the database and not taken from the text of Thoreau's journal entry.

Each of the botanical notes was then classified, on the basis of context, into a number of conceptual categories, including location/landmark; description; phaenological stage; habitat; ethological interaction; or usage. Sorting the database material by category and by species provided locational information for many plants, the precise whereabouts of which had not been consistently recorded in the journal. Sorting along other category combinations allowed for further refinement of location and the resolution of many ambiguities in the recording of geographic or environmental data associated with a number of the plants.

A final matrix column was reserved for the text of the journal entry about each plant, in complete, or more often abbreviated, form. This "narrative" area was designed to provide an extensive data space in which long verbal descriptions could be recorded and parsed. Here, for example, important information about the specific landmarks and boundaries of an area in the vicinity of the landmark plant might be kept for further analysis. For example, a textual passage containing the fact that walking around Fair Haven Hill from the Cliffs side past the west spring woods brought one to the boundary of Abiel Wheeler's field might prove important in determining the location of a plant which later might be identified in that field - and about which no other information might be forthcoming in the field note written to describe it. In this column, also, data on lines of sight from a number of plant-related locations could be entered, contributing to the creation of linkages between bits of text about otherwise unrelated individuals (for example, the fact that from the top of Fair Haven Hill in Concord, one could look across to the aspens behind James Baker's home in Lincoln).

Elm (Slippery)	Ulmus rubra (SUPPERY ELM)	05.	312	4-Jul	1853	7/4/1853		> Lee's Cliff via Comer Spnng	Lee's Cliff	location	at Lee's Cliff, under the slippery elm, Panetana P
Parietaria pensylvanica	Panetana pensylvanica (COMMON I	05.	312	4-Jul	1853	7/4/1853	N	> Lee's Cliff via Comer Spring	Lee's Cliff	order: flowenng	at Lee's Cliff, under the slippery elm, Panetaria P
Pellitory (American)	Parietana pensylvanica (COMMON I	05.	312	4-Jul	1853	7/4/1853	N	> Lee's Cliff via Comer Spring	Lee's Cliff	order: flowenng	at Lee's Cliff, under the slippery elm, Panetaria P
Anychia dichotoma	Paronychia canadensis (FORKED C	05.	312	4-Jul	1853	7/4/1853	N	> Lee's Cliff via Comer Spnng	Lee's Cliff	order: flowering	at Lee's Cliff, under the slippery elm, Panetaria P
Chickweed (Forked)	Paronychia canadensis (FORKED C	05.	312	4-Jul	1853	7/4/1853	N	> Lee's Cliff via Comer Spnng	Lee's Cliff	order: flowenng	at Lee's Cliff, under the slippery elm, Panetaria P

Fig. 3. A segment of the data matrix, showing the elements of the sentence "At Lee's Cliff, under the slippery elm, Parietaria Pensylvanica, American pellitory, in flower, and nearby Anychia dichotoma, forked chickweed (Queria) also in flower ," from Thoreau's walk to Lee's Cliff via the Corner Spring, July 4, 1853. The slippery elm is used in this set of entries as as a landmark/location indicator. The first column of the matrix contains the common name of the plant; the second, the scientific name; the fourth, the page in the printed edition of Thoreau's journal; the fifth, sixth and seventh, the date of the entry. In the 8th column, "N" indicates that alternative nomenclature is used for the plant in the passage analyzed. Column 9 contains route information; column 10, the exact location of the plant; Column 11, the conceptual category; and Column 12, the full text of the journal entry.

The information in the data matrix, taken as a whole, might then be used to serve in the process of "districting," or the creation of superordinate regions comprising a number of smaller, discrete subdivisions of space, each of which has a specific name of its own. The fact that Thoreau occasionally did this in his own notes, as, for example, when he created the concept of the "Esterbrooks Country" to contain fifteen or twenty subordinate, individually named regions, each of which he might previously have thought of as separate spaces or habitats, led the experimenter to posit some other large-scale, inclusive "environs" within the data space, which, though perhaps not named as such by Thoreau, have the advantage of subsuming under one inclusive name a host of small, spatially-contiguous individual habitats which, if sorted alphabetically in the database by their individual names, might be difficult to conceptualize in their actual, organized relationships to one another on the ground.

Although it had at first seemed desirableto keep the database free of information derived from sources other than Thoreau's journal, this superordinate districting process was eventually supplemented by the occasional use of a street name derived from other contemporaneous or modern sources. This helped to organize the location of house-centered habitats mentioned in the pages of the journal. For example, "Minot Pratt's" is not explicitly stated by Thoreau to have been located on Monument Road in Concord, but adding that information allows the database user to position it in the vicinity of other houses also located along Monument Road. In his work as a surveyor, Thoreau consciously used appropriate street names in the drawing of lot lines for his clients, so the importation of this information from ancillary sources was not thought to violate the integrity of the database material, since one of the points of the exercise was, in so far as possible, to pinpoint the location of natural objects which Thoreau had identified and observed, and to chart, if possible, their spatial relationships to one another. Thoreau did not omit street names in order to conceal his ecological finds, but simply because such information was so obvious as to go, quite literally, without saying in a set of notes he kept for his own personal reference.

While far from complete at the present time, the data matrix was conceived to function as an instrument for the resolution of ambiguity and missing information in verbal and textual descriptions of a specific space. In addition, it was hoped that the processes of sorting and refining text-based geographic information would allow the user to project, on the basis of words alone, a reasonably accurate map of the space described - to assemble, from the immense verbal jigsaw puzzle of Thoreau's ecological notes, a comprehensive and unified picture of the environment he so carefully documented over the course of the last twelve years of his life. To date, this map has not yet been projected, although this would be the logical next step in the project as it has progressed to the present time.

Several maps of "Thoreau's Country" already exist, and the creator of one of them, Herbert Gleason, had the advantage of calling upon the recollections of individuals still alive at the time (1902) who had walked with Thoreau or knew the local place names to which he referred. But none of them is as complete as a fully-dedicated use of Thoreau's text might permit them to be, and each of them suffers from a number of inaccuracies, at least some of which might be easily correctable by close reference to Thoreau's text.[6] It should be noted, however, that the language of Thoreau's spatial references, while perfectly accurate, is not always without ambiguity; there are passages in his notes which cannot be plotted properly on a map unless alternative parsing protocols are used. [7] These passages might provide interesting test material to demonstrate an individual's ability to create a graphical representation from spatual information presented verbally.

Like the terrain exploration itself, the creation of this database proved to be more labor-intensive than might be desired, but it was hoped that the process of creating it might serve as an object lesson in the development of an instrument which might automate the process of parsing and resolving discontinuous bits of verbally-based

6See, for example, the mislocation of the kalmia swamp and Beck Stow's swamp in the most recently published of these endeavours, on the endpapers of Bradley Dean's 1999 edition of Thoreau's last known manuscript, "Wild Fruits."

7Perhaps the best example of this is the reference to the little white pines growing beyond Abiel Wheeler's house on the Corner Road. The author posits that it is virtually impossible to project an accurate representation of the spatial relations amongst theseparateelements named in this passage of Thoreau's journal (November 8, 1860) without an understanding of the way he used the elements of spatial language.

geographic information, in such a way that it might be used as the foundation of a two- or three-dimensional map of a specific space, and ultimately, a computer-mediated model of an actual environment. The Musketaquid ground is unique in having been so painstakingly documented by a man whose interest in spatial relations was subordinated only to his interest in the relations amongst living beings sharing and inhabiting the space, but the same techniques which allow for the pinpointing of places within that six-mile square should also be useful in analyzing, learning, and comprehending other spaces, as described in the words of people who know them well.

Two additional dimensions of the matrix might be mentioned here, although they have no immediate relation to the geographic and locational functions it was designed to serve. When the database material is has reached a its ultimate stage of completeness and accuracy, sorting the matrix on the category of "location" will produce a "botanical habitat list" detailing the entire range of plants which Thoreau found in each of the many individual environments he chose for intensive study. This might be of interest to field ecologists and historians of land use in the local area, as well as to students of Thoreau's methods as a natural historian. In addition, sorting by the "date of entry" will produce a phaenological calendar for the entire region, indicating the days on which the individual plants Thoreau kept under observation entered the various stages of their annual life cycles. There is evidence that before he succumbed at the age of 44 to his lifelong battle with tuberculosis, Thoreau was organizing the material in his thousands of pages of journal notes - estimated by some scholars to total two million words - to produce a descriptive "Kalendar" of Concord, detailing the unfolding of a single ideal year in the natural history of the region he so loved. This monumental undertaking, left unfinished in the form of lists, charts, tables, and hundreds of pages of barely legible notes at the time of his death, would have been made much easier had he enjoyed the use of the electronic databasing tools available to a modern researcher.

Future plans for this project include the development of a map of the area, using GIS tools available through the spatial information project of MIT's Information Systems department, and an analysis of Thoreau's original land surveys, currently being processed for access over the internet by the staff of the Concord Public Library. An article on Thoreau's hitherto unrecognized work in the cognitive area of spatial information processing is also in preparation, to demonstrate that what was described as an"instinct for locality" was in fact a learned facility with spatial relations that reached the apparent effortlessness of artistry only through years of patient, painstaking and continuous effort with map and compass, research and experiment, walking with open eyes in the way that any thinking individual can, with like practice, also learn to do.

Acknowledgments

The experimenter wishes to thank the following individuals for encouragement and assistance during the course of this experiment: Peter Gifford and John Nickols, lifelong residents of Concord, whose spontaneous conversations with an unknown walker in their countryside led to an extension of horizons, both material and immaterial; Marilyn Lawrence of the Concord Natural Resources Commission, Ann Wanzer of Concord's "Books with a Past," and Thomas Gumbart of the Lincoln Conservation Commission, for their enthusiasm about this work, their insights and their respective offerings of excellent and otherwise unattainable maps; Peter Weld, Dorrie Hall, Donna Neal, and Lynn Dupont Wood, who were kind enough to participate vicariously - and even on occasion actually - in what less sensitive individuals might have been prone to dismiss as a persevering and insurmountable obsession; Rudy Darken, of the Naval Postgraduate School, whose quote from James Fenimore Cooper in the guest editors' introduction to the first special issue of *PRESENCE: Teleoperators and Virtual Environments* on Spatial Orientation and Navigation in Large-Scale Virtual Spaces (7.2, February 1998) sparked a discussion of nineteenth-century literary ancestors of the science of spatial knowledge acquisition, of whom Thoreau seems certainly to have been one; Kari Anne Kjolaas, of MIT's Research Laboratory of Electronics, whose interest in the idea of "Virtual Walden" sustained the experimenter in those early days when it would have been easy to let the concept die of its own immense and ever-increasing weight; and finally Nat Durlach of MIT's Research Laboratory of Electronics, Principal Investigator of the VSPAN project for the Office of Naval Research, whose budgetary shortfall came at exactly the right time to give the experimenter the requisite unencumbered leisure in which to pursue the Hermit of Walden on his path through the fields and woodlands of Musketaquid - the venue of his own private experiments in the synthesis of a space-based, imaginal world.

References

The section headings of this paper are from the following locations in Thoreau's writings:

Walking may be a science, so far as the direction of a walk is concerned.
Journal, August 22, 1854

What is a course of history or philosophy, or poetry, no matter how well selected, or the best society, or the most admirable routine of life, compared with the discipline of looking always at what is to be seen?
Walden (1854)

Do not take a dozen steps which you could not with tolerable accuracy protract on a chart. I never do otherwise.
Journal, February 15, 1857

Could a greater miracle take place than for us to look through each other's eyes for an instant?
Walden (1854)

How circumscribed are all our walks, after all! With the utmost industry we cannot expect to know well an area more than six miles square ...
Journal, March 18, 1858

The *Botanical Index to the Journal of Henry David Thoreau* is available as a publication of the Thoreau Quarterly (1983); as volume 15 of the Peregrine Smith reprint of the Houghton Mifflin 1906 Riverside Edition of Thoreau's Journal, and on the web, through the support of the Harvard University Herbaria, at http://www.herbaria.harvard.edu/~rangelo/BotIndex/

Sources consulted for the specific topography and environmental regions of the area under study are as follows:

Eaton, R. (1974). *A Flora of Concord*. Cambridge: Museum of Comparative Zoology.

Lincoln Land Conservation Trust: (1992). *A Guide to Conservation Land in Lincoln*.

Stowell, R., and Howarth, W. (1970). *A Thoreau Gazetteer*. Princeton: Princeton University Press.

Thoreau, H.D. (1906). *Journal*. 14 volumes. Boston: Houghton Mifflin.

Sources which were especially helpful in learning the technical process of wayfinding are:

Department of the Army (1969). *Survival, Evasion and Escape*. Field Manual #21-76. Washington, D.C.

___________ (1987). *Map Reading and Land Navigation*. Field Manual #21-26. Washington, D.C.

Paul, D. (1991). *The Green Beret's Compass Course*, 6th edition. Woodland, CA: Path Finder Publications.

Pepi, D. (1985). *Thoreau's Method, A Handbook for Nature Study*. Englewood Cliffs, NJ: Prentice Hall.

APPENDIX: The Random Walks

To my astonishment I was informed on leaving college that I had studied navigation!—why, if I had taken one turn down the harbor I should have known more about it..

- Henry Thoreau, *Walden*

(Year 1, 1997)

5/30	Concord Village: Main Street, Turnpike and Boston Roads
6/6	North Bridge via Monument Road and thence to Walden Pond
6/20	Concord Village: Turnpike
7/11	Concord Village: Library, Bedford Road, Sleepy Hollow
7/24	Sleepy Hollow and Walden Pond
7/31	Walden Pond Concord Museum (to look at historical maps)
8/8	Great Meadows
8/27	North Bridge, Battlefield, and Buttrick's farm
9/4	Walden Pond, Emerson's Cliff
10/7	Concord Village, Milldam, Sleepy Hollow
10/29	Esterbrooks Country via Old Carlisle Road
11/14	Nawshawtuct; thence to Nine Acre Corner via old Sudbury Road and return over Conantum
11/28	Fair Haven woods and Mount Misery, Lincoln
12/4	Concord Village, Turnpike (1/2 day)
12/26	Walden woods and Heywood's Meadow
12/31	Great Meadows via Bedford and Virginia Roads Nawshawtuct

Year 2 (1998)

1/19	Dugan Desert via Corner Road and C. Miles'; return by Marlborough Road

1/26	North Bridge via Merrick's Pasture Turnpike past Emerson's to Hosmer's farm
2/13	Hosmer's Pond and Brooks' Corner, Lincoln; return via Emerson's
2/16	Brister's Hill, Hubbard's Close and Walden Pond
3/16	Barnes Hill via Dakin's Brook [Lowell Road] and return via Buttrick's
3/17	Lee's Bridge via Corner Road and C. Miles', return over Conantum
3/18	Fair Haven Cliffs Nawshawtuct
3/19	Lane between Hawthorne's and Tuttle's (1/2 day)
4/15	Esterbrooks Country and Punkawtasset Hill; return via North Bridge
5/15	Esterbrooks Country, Punkawtasset Hill and Yellow Birch Swamp
5/21	Baker Farm [Lincoln], Heywood's Brook and Walden woods
6/1	Lee's Bridge via Marlborough Road and Powder Mill Road; return through Mount Misery, Heywood's pasture, Walden woods and Brister's Hill
6/5	Esterbrooks Country via North Bridge; Boulder Field and Yellow Birch Swamp; return through Sleepy Hollow
6/9	Great Meadows
6/11	Flint's Pond [Lincoln]
6/29	Lee's Bridge via Conantum and back by bridge below Clamshell
7/2	Fair Haven Cliffs via Hubbard's meadow; return via Walden Pond
7/23	Mill Brook environs (1/2 day)
8/6	Ball's Hill via Monument Road
8/16	Pleasant Meadow via Codman woods [Lincoln]; return via Corner Road, C. Miles' and Nut Meadow Brook
8/19	Nawshawtuct and Mill Brook (1/2 day)
8/21	Riverside at Merrick's pasture and thence on Lowell Road past J. Farmer's
8/27	Assabet One Arch Bridge and Union Turnpike to Rotary and return via

	Barrett's Mill
	Saw Mill Brook at Hosmer's Farm
9/15	Ministerial Swamp via Marlborough Road; thence toWhite Pond; return via Baker Farm, Walden and Brister's Hill
9/16	Punkawtasset via North Bridge
9/17	Gowing's Corner via Emerson's; return through Great Meadows
9/18	Hathaway Hill [Drumlin Farm], Lincoln
9/29	Fair Haven Cliffs via Baker Farm
10/16	Flint's Pond and Pine Hill [Lincoln]
10/21	Mount Misery; return via Corner, Powder Mill Road, Second Division and Ministerial Swamp
10/23	Buttrick's farm; thence to Barrett's Assabet meadow
11/6	Codman woods, Baker Farm, Andromeda Ponds, and Heywood's Peak, Walden
11/12	Esterbrooks Country: Lime Quarry and Boulder Field
11/19	Hubbard's Bridge (1/2 day)
11/20	Mackintosh's swamp, Pine Hill and Goose Pond [Lincoln], Brister's and Hosmer's farm; return along Turnpike
	Assabet River over Nawshawtuct
11/27	Annursnack Hill and fields westward [Acton]
12/10	Brister's Hill and Hubbard's Close (1/2 day)
12/15	Hubbard's meadow, Miles Road, Dugan Desert, Tarbell's and E. Wood's
	Boston Road: Gowing's corner and ridge behind Alcott house
12/28	Old Bedford Road via Bedford Road and back via Boston Road (1/2 day)

Year 3 (1999)

1/12	Hubbard's Bridge and Mill Brook crossing (1/2 day)
1/21	Walden Pond via Walden [poorhouse] road (1/2 day)
1/29	Union Turnpike via E. Wood's 91/2 day)

2/4	Esterbrooks Country
2/10	Sudbury Bound via Codman's Swamp, Pine Hill, Farrar's, and Lee's Bridge [Lincoln]; return by Conantum
2/12	Boston Sportsmen's Show (to buy maps and look at GPS devices) (1/2 day)
2/15	Wier Hill and Pantry Brook, Sudbury, from Lincoln, via Sherman's Bridge Wayland] and back via Corner Spring
2/19	Clematis Pond via Mt. Misery, and return via Well Meadow and Fair Haven Cliffs Mill Brook above Emerson's
3/5	Assabet around Nawshawtuct Concord Library: read "Thoreau mapping project" in archive collection
3/10	Ministerial Swamp via Dennis' swamp, Damon's factory and Tarbell's farm
3/17	Concord Village, Mill Dam (1/2 day)
3/29	Codman Swamp and Pine Hill [Lincoln] (1/2 day)
3/30	C. Miles Swamp; thence to Nine Acre Corner and back via Conantum
3/31	Esterbrooks County and to Carlisle via Old Carlisle Road
4/2	Cedar Hill [Lincoln] via Goose Pond and Saw Mill Brook; around Flint's Pond; return over Bear Hill and Walden Road
4/8	Back Road and Deep Cut Woods to Heywood's Brook and Mt. Misery; return via Well Meadow and Heywood's Meadow
4/16	Sleepy Hollow and by Peter's Path to Great Fields (1/2 day)
4/19	Carlisle Bridge [Carlisle] via Buttrick's; return via Abner Buttrick's Hill and Monument Road
4/21	Along "Battle Road" from Lexington through Lincoln to Concord
4/27	Flint's Pond via Wheeler Farm, and return via Three Friend's path [Lincoln]
5/6	Lincoln: Silver Hill bog and Flint's farm
5/11	Fair Haven Cliffs, Mount Misery and Baker Farm
5/21	Bedford Road to Bedford, returning through Copan, Great Meadows and

Moore's swamp

5/27	Beaver Pond, Lincoln, via Weston bound on Great Road; return over Pigeon Hill and Coburn Bog
6/2	Assabet River via Wheeler's meadow and Squaw Sachem's trail (1/2 day)
6/3	Smith's Hill, Lincoln, via Turnpike and Brooks Farm; return via Elm Brook and Boston Road
6/11	Esterbrooks Country: Two-Rod Road via Monument Road and Punkawtasset River at Hurd's Bridge
6/15	Harrington's, Second Division Brook and Loring's Pond Conantum via Tarbell's and C. Miles'
6/18	Mount Tabor and Flint's Pond [Lincoln]
6/24	on river to bend above Sherman's Bridge
7/8	Hubbard's meadow Barrett's meadow via Hunt's Bridge and return via Buttrick's shore
7/15	Esterbrooks Country: Mill site, Hubbard's Hill and Pratt's brook
7/21	Well Meadow via Baker Farm; return on Fair Haven back road River and south base of Nawshawtuct
7/30	Ministerial Swamp via Tarbell's
8/5	Sudbury via Old Marlborough Road; return through Corner and Mount Misery
8/20	Sleepy Hollow via Poplar Hill path (1/2 day)
9/1	Tanner's Brook [Lincoln] via Boston Road; thence to Lincoln Hill
8/20	Concord Village; river at Hurd's Bridge (1/2 day)
10/6	Poplar Hill (1/2 day)
10/7	Brister's Hill; return via Hubbard's Close and Turnpike (1/2 day)
10/8	Hubbard's meadow and bridge (1/2 day)
10/11	Harrington's and Ministerial Swamp; thence to White Pond via C. Miles and Corner; return via Powder Mill and Old Marlborough Road
10/12	Wheeler's meadow, Nawshawtuct Common and Woodis Park (1/2 day)

10/13	Fair Haven woods via Back Road; return via Pine Hill
10/15	Down Boston Road (1/2 day)
10/19	Acton via Dodge's Brook, Barrett's Mill and Annursnack, return via Wetherbee's Farm, Lowell Road, river at French's rock, and North Bridge maple
10/21	Pine Hill via Goose Pond, thence to Saw Mill Brook and return via Walden Pond
10/28	Mill Brook (1/2 day)
11/2	Farmer's Cliff and Bateman Pond woods; return via Barnes Hill and North Bridge maple (cut down)
11/5	Assabet hemlock plain via stone bridge (1/2 day)
11/9	Ed. Hosmer's farm, Lincoln (1/2 day)
11/23	Maynard's [Sudbury] via Codman woods and Farrar's [Lincoln]
12/8	Austin's, Flint's Pond and Pine Hill [Lincoln] (1/2 day)
12/16	Concord Village, Library (1/4 day)
12/23	Hurd's Bridge and Nawshawtuct meadow (1/4 day)
12/24	Assabet Stone Bridge via Colburn Hill (1/2 day)

Year 4 (2000)

1/3	Esterbrooks Country: Hubbard's Hill and Yellow Birch Swamp via Hunt's Pasture
2/3	Great Meadows via Boston Road and Old Bedford Road; return via Great Fields
2/10	Nine Acre Corner via Dennis' swamp and Sudbury Road; return over Conantum
2/23	Walden woods via Baker Farm and Heywood's Brook Riverside at Merrick's Pasture

TRAINING TRANSFER EXERCISES:

8/15/97	World's End Reservation, Hingham
10/8/87	Salem
10/13/97	World's End Reservation, Hingham
11/7/97	Salem
5/12/98	World's End Reservation, Hingham
7/9/98	Salem
7/20/98	Salem
7/21/98	Salem
8/13/98	Rockport
9/1/98	World's End Reservation, Hingham
9/4/98	Rockport
9/25/98	Salem
10/29/98	Salem
3/18/99	Quincy Historical Society (to look at maps)
4/9/99	Gloucester
6/4/99	Boston Public Gardens and Common
6/14/99	Boston: Arnold Arboretum
9/24/99	Rockport
12/17	Salem

TERRAIN OBSERVATION EXCURSIONS:

7/18/97	by train to Fitchburg
9/27/97	by car to Carver via Plymouth
10/31/97	Hancock Observatory (for elevated view of terrain between Boston and Concord)
1/13/98	by car to Carver via Plymouth
4/20/98	by car to Cape Cod
6/7/98	by car to Cape Cod
10/5/98	by car to Nashua
10/10/98	by car to Carver via Plymouth
11/11/98	by car to Saugus
11/14/98	by car to Attleborough
12/16/98	by train to Newburyport
12/18/98	by train to Plymouth
12/23/98	by train to Haverhill
12/24/98	by train to Fitchburg
1/5/99	by train to Lowell
2/6/99	by train and car to Dudley Pond, Wayland
2/8/99	by train to Worcester
2/9/99	by train to Worcester
4/29/99	by train to Fitchburg
8/12/99	by car to Cape Cod
10/14/99	by train to Worcester
10/28/99	by train to Fitchburg
11/10/99	by train to Canton
11/12/99	by car to Saugus

Inner Space: A Rat's Eye View

Andrew Brooks

Research Laboratory of Electronics
Massachusetts Institute of Technology

Introduction

This document is intended as a brief companion piece to ``A Random Walk: Mastering Musketaquid'', from a team member who has a personal interest in the acquisition of spatial knowledge (in both configurational and route forms) of complex man-made three-dimensional interior spaces, as opposed to the natural spaces described in the aforementioned treatise. It is intended in a similarly introspective, anecdotal sense. The author is interested in a form of architectural exploration known as ``urban speleology'', or less formally as ``infiltration'' and, on certain college campuses, as ``hacking''. A variant that makes primary use of storm drains as a traversal mechanism is known colloquially as ``draining''. The major goal of this hobby is the discovery of unusual, interesting or forgotten spaces within buildings or other human constructions. Other aims are sometimes to determine unusual routes of passage between common, easily visited areas, and methods of travel to places that are not easily accessed by conventional means.

Background

When major edifices such as buildings are constructed, design proceeds in layers. First the superstructure is put in place, then all the subnetworks such as electrical, plumbing, and air conditioning ductwork. In addition, the lifetime of most buildings is very long, so their interior workings usually undergo many changes during this time. Furthermore, each layer of design is generally more concerned with cost and the ability to perform maintenance, rather than universal accessibility and optimum use of space, which are largely conflicting requirements in any case. The combination of these factors results in most buildings that have been established for some time having interiors that are very three-dimensionally complex and contain a certain quantity of space that is neither occupied by solid physical objects nor easily accessible. These spaces, when big enough to warrant consideration, are known as ``tombs'' and are the grails of urban speleologists. Discovery of such tombs requires a keen eye for spatial observation and finely honed spatial navigation and orientation skills.

A typical exploratory expedition involves the selection of a building or site that looks promising, and ``scouting'' the area from the outside and from common areas such as hallways and restrooms, to get an idea of the shape and layout of the building and its levels. Often floorplans will also be consulted if they are available. By definition floorplans rarely show forgotten spaces at all, and are usually out of date, but one can often use them to identify promising starting points and to gain more of a coarse spatial awareness of the two-dimensional arrangement of each floor. However, despite usually being helpful to some degree, two-dimensional maps such as floorplans can never have fine enough detail for urban speleology. There are too many different types of access point, from hatches in walls and ceilings to air conditioning ducts to random holes in interior structures long since hidden from view and ignored. In addition, these interconnections typically involve enough detail in the third dimension that it would be difficult if not impossible to effectively render this information in two-dimensional form. The actual searching for tombs must occur in person inside the environment.

Spatial Aptitude Factors

Foraging for spaces consists of two main endeavours: looking for visual cues, and physically following entryways to see where they lead. In terms of visual cues, one must train oneself to notice spatial arrangements that are out of the ordinary; where areas of accessible space or physical barriers do not fit together evenly, possibly indicating ``missing'' space. One of the most important spatial skills seems to be the ability to judge relative lengths and volumes accurately, when only one of the quantities is in view at a given time. Only if this skill is well developed will it be possible to make competent hypotheses as to the space consumption of accessible spaces using purely visual information. It is also important to be able to think of the location in more than just two-dimensional slices, such as floors. The desired spaces may lie between these slices, or may not fit easily into a floorplan mentality; the spaces are frequently small enough such that they may reside on a given floor but easily allow the explorer to pass over or underneath them without noticing their existence.

The act of physically exploring entry points within these complex, cramped environments exposes spatial navigation factors that are to some extent separable into two categories; localisation and orientation. Localisation is a matter of determining where, in terms of three-dimensional spatial co-ordinates, the explorer is located within the structure. Orientation is determining where the explorer is facing, in terms of some principal axes of the structure or the external environment. Although these can be separated to a degree, a very important function that involves both classes is the ability to determine which space or area lies ahead, behind, to either side, above and below the explorer, even though most or even all of these directions are likely to be visually occluded. This is consequently the most difficult spatial task in urban speleology. Experience of this researcher has shown that performance in this task is able to be enhanced by qualitative improvements in the localisation and orientation subtasks, so specific comments on these will be presented.

Man-made structures, while complex, usually exhibit some degree of regularity in their spatial arrangement. A principal problem with localisation therefore is the ``junction problem''; how to cope with the intersection of areas having a significantly different spatial arrangement. This problem occurs both in the physical environment and when studying floorplans, as it does to a lesser degree with maps of natural environments.

Unlike maps of natural environments, in which the terrain along the junction may usually be estimated with smooth mental interpolation to a reasonable degree of accuracy, junctions of representations of man-made spaces and the spaces themselves tend to be very abrupt and often do not link in any particularly intuitive fashion. A simple example is the case when two buildings with different floor heights are abutted in such a way as to offer foot passage between their hallways. Some hallways in one building will interconnect with hallways in the other building, perhaps involving a small amount of elevation correction such as a ramp or set of stairs, while other hallways will encounter dead ends. This can be very confusing to an individual that is unfamiliar with the spatial arrangement of one or both of the buildings.

The researcher has reached two conclusions based on experience with the junction problem. Firstly, the only effective way to acquire spatial knowledge about the junction is to experience it first hand. Floorplans are of little use, for they cannot contain the level of detail necessary to effectively visualise the transition, even if they contain extra spatial information such as floor elevations. Secondly, redundant cues in the environment can greatly help with localisation, both across junctions and in general. For example, many campus buildings are catalogued with a consistent numbering scheme -- the building itself is assigned a number, as are the floors of the building and all well-defined rooms, with all rooms on a particular floor commencing with the number of the floor. In addition, some buildings incorporate colour-coding of walls and other structures that can identify the general position of a space within the building's strata. This is by no means a complete solution to the junction problem -- what to do with a room that spans two floors, for example -- but can provide a helpful frame of reference for an explorer, for whom self-localisation is then assisted by relative translation from a nearby known and named space. This coarse granularity, similar to a spatial mnemonic, is thus useful for making the transfer from an extremely coarse spatial representation, such as an external map, to the very fine mental representation of the space.

The case of orientation may seem simpler than localisation, at least in the terrestrial environments that this researcher has had an opportunity to explore, since `up' and `down' are always accurately represented due to gravity. Nonetheless, maintaining directional information within an enclosed space feels more difficult than estimating one's translational position. This may be due to the ability to ``chunk'' spaces into rough three-dimensional localised arrangements; for example, it is reasonably easy in general to keep track of whether one is vaguely above, below, to the left or to the right of a previously visited space. Performing this approximation with directions seems more complicated within a complex indoor environment. Perhaps the restricted field of view associated with small enclosed spaces makes judging angles of self-rotation less accurate; this would cause one's heading error to become large rapidly when navigating through direction changes that are not integer multiples of ninety degrees. Layouts not conforming to a rectangular grid are frequently encountered within the confines of buildings' viscera.

To some degree, the ability to recognise external boundaries of the environment (in the case of a building, these would be the exterior walls) from the inside can provide a powerful source of correction of orientation error. However, due to the typically short range of visibility, in general external boundaries can only be recognised when one is physically close to them. Other than this, few specific cues are of much assistance. If the path being followed takes one through multiple buildings or named surroundings, however, a coarse external map can be of use in keeping one's sense of direction narrowed to within a rough range, such as about the major and minor compass points.

It is also worth noting that the orientation category also suffers from the junction problem. Consider the case of moving from one straight subterranean tunnel to another via a circular manhole. It is difficult to view the axes of both tunnels simultaneously, and the rotationally invariant shape of the manhole gives no directional reference cues from one tunnel to the other. Unless the tunnels are aligned to be parallel or perpendicular, it is likely that a significant bearing error will be introduced by the transition. Besides the careful physical placement of the body during the transition, little can be done to maintain an accurate spatial representation short of manufacturing additional orientation cues at the junction.

Further Discussion

Urban speleology presents problems for formal experimentation, both in the real and virtual cases. In the real world, finding a suitable location for experimentation could prove difficult, as could recruiting subjects with the required willingness and experience to put themselves through an arduous, dirty and dangerous environment not explicitly designed for human navigation. In the virtual domain, the level of three-dimensional complexity makes modeling of the environment, already difficult even for relatively straightforward spaces, a challenging prospect indeed. In addition, the traversal of these hidden spaces in the real world is a physically intense process, which may distract the explorer from constructing a mental model of spatial awareness, but also may emphasise certain aspects of the resultant representation. The inability to simulate this level of concentration on physical movement and exertion within the virtual environment may detract from the effectiveness of extrapolating performance in this mode to that in the real world, and might also impact studies of training transfer.